

[illegible]

[illegible]

Genature version 5.1.3
Copyright (c) 1993 - 2003 Computer Ltd.

om nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:20:22 : Search time 2934.92 seconds
(without alignments)
18469.506 Million cell updates/sec

Index: us-09-026-459a-42

Index score: 6447

Sequence: 1 GGGGTATGGGAGGAAAC.....AAATGAGGATTAATGATAGT 3347

Scoring table: IDENTITY 90%

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 804774376 residues

Total number of hits satisfying chosen parameters: 32498142

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match ok

Maximum Match 100%

Listing first 45 summaries

Database :	EST *
1:	em_estb01 *
2:	em_estb02 *
3:	em_estb03 *
4:	em_estb04 *
5:	em_estb05 *
6:	em_estb06 *
7:	em_estb07 *
8:	em_estb08 *
9:	em_estb09 *
10:	em_estb10 *
11:	em_estb11 *
12:	em_estb12 *
13:	em_estb13 *
14:	em_estb14 *
15:	em_estb15 *
16:	em_estb16 *
17:	em_estb17 *
18:	em_estb18 *
19:	em_estb19 *
20:	em_estb20 *
21:	em_estb21 *
22:	em_estb22 *
23:	em_estb23 *
24:	em_estb24 *
25:	em_estb25 *
26:	em_estb26 *
27:	em_estb27 *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	910.2	27.2	1624	13	BM459031 AGENCOURT
2	909.4	27.2	1051	14	BM561083 AGENCOURT
3	856	25.6	1083	14	BM439221 AGENCOURT
4	844.8	25.2	881	14	BM220275 AGENCOURT
5	762.4	22.5	1122	13	BM545881 AGENCOURT
6	725.4	21.7	727	9	AL597811 DKFZP3130

7	699.6	20.9	824	12	BM610661
8	698.2	20.9	792	9	AL563757
9	692.2	20.7	767	12	BM579456
10	691.8	20.7	922	13	BM466194
11	678	20.3	835	13	BM453724
12	655.2	19.9	1001	13	BM475604
13	660	19.7	672	13	BM186859
14	646.6	19.3	882	14	BM222227
15	643.6	19.2	797	10	BM539278
16	642.2	19.2	581	12	BM516239
17	617	18.4	618	13	BM264009
18	614	18.3	694	10	BM082846
19	605.2	18.1	623	10	BM383181
20	605	18.1	605	9	AL599633
21	596.4	17.8	922	12	BM253543
22	580	17.3	580	9	AL599105
23	576.4	17.2	1364	11	AK011246
24	575.2	17.2	914	9	AA763411
25	566.4	16.9	588	10	AA762234
26	558	16.7	871	9	AA763485
27	530.2	15.8	598	10	AA762967
28	530	15.8	539	9	AL692790
29	528	15.8	572	9	AL130943
30	516.8	15.4	597	12	BM311246
31	513.2	15.3	594	10	AV715533
32	512.2	15.3	734	12	BM144857
33	505.2	15.1	736	13	BM151568
34	493.8	14.8	525	9	AA258255
35	486.4	14.5	584	12	BM149050
36	484.4	14.5	493	10	AA502887
37	481.4	14.4	484	10	BM168095
38	480.2	14.3	659	13	BM211116
39	476.6	14.2	2083	13	BM456728
40	474.4	14.2	734	13	BM152682
41	459.2	13.7	497	12	BM347844
42	457	13.7	457	9	AI094215
43	457	13.7	742	9	AL598766
44	454.2	13.6	869	12	BM65749
45	447	13.4	447	9	AI082179

ALIGNMENTS

RESULT 1
BM450031
LOCUS BM450031 1023 bp mpna linear EST 05 FEB 2002
DEFINITION AGENCOURT_6393544 NIH_M50_72 Homo sapiens cDNA clone IMAGE:5528037
5' mRNA sequence.
ACCESSION BM450031
VERSION BM450031.1 GI:18499071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1023)
NIH-MGC <http://mgc.nhlbi.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cuapbs@mail.nih.gov
Tissue Procurement: AACC/WU/DHP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: IAM2204 row: d column: 22
High quality sequence stop: 643.
Location/Qualifiers
1..1023

Plate: LLAM8427 row: 1 column: 18
High quality sequence stop for

FEATURES

Source

Location: GenBank

1..797

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3449873"

/clone_lib="NIH_MGC_10"

/cell_line="MGC36"

/lab_host="DH10B"

/note="Organ: corvix; Vector: pCMV-Sport6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.5 kb. Library prepared by Life

Technologies."

BASE COUNT 252 a 171 c 167 g 207 t

*K161N

Query Match 19.2% Score 643.6; DB 10; Length 797;

Best Local Similarity 95.8% Pred. No. 2.2e-115;

Matches 747; Conservative 0; Mismatches 24; Indels 9; Gaps 8;

QY 2000 CUGGTCAGGACATTCACAACTCTTTCATCAACAGACAGACGATATGATTCATTATTATAG 2059

DE 1 CTGTTTCAGGACATTCACAACTCTTTCATCAACAGACAGACGATATGATTCATTATTATAG 59

QY 2060 TATTCATTAACCTGCTTTCATGAGAGACTGAAACAAATATTTTGCAGTATGCTTCCA 2119

DE 60 TATTCATTAACCTGCTTTCATGAGAGACTGAAACAAATATTTTGCAGTATGCTTCCA 118

QY 2120 GAGAGGCGGCTATCTTGTCAGCAATAGCTACATTCCTGCGAAGGCGCTTACAGTTTCTTA 2179

DE 119 CTAGGCGGCTATCTTGTCAGCAATAGCTACATTCCTGCGAAGGCGCTTACAGTTTCTTA 178

QY 2180 GTTCATGCTTACGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 2239

DE 179 GTTCATGCTTACGCAATTCCTGAGGGAACATCTATATTCACCCGCTGAGAGTCCATATA 238

QY 2240 AAATTTCAGAGCTTTCAGCAATAGCTACATTCCTGCGAAGGCGCTTACAGTTTCTTA 2299

DE 239 AAATTTCAGAGCTTTCAGCAATAGCTACATTCCTGCGAAGGCGCTTACAGTTTCTTA 298

QY 2300 CAATTGCTGAATCATTCGCGGATCTCTGCAAGATTCAGCAATTAATGATGATGCTA 2359

DE 299 CAATTGCTGAATCATTCGCGGATCTCTGCAAGATTCAGCAATTAATGATGATGCTA 358

QY 2360 AATGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 2418

DE 359 AATGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 418

QY 2419 AAATCAGCTTTTATATTCAGGATTCAGCAATTAATGATGATGCTA 2478

DE 419 AAATCAGCTTTTATATTCAGGATTCAGCAATTAATGATGATGCTA 478

QY 2479 GATTCGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 2537

DE 479 GATTCGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 538

QY 2538 GATTCGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 2597

DE 539 GATTCGCAATTCCTGAGGGAACATCTATATTCACCCGGAAGAGTCCATATA 598

QY 2598 CTTTGGTGGATTCAGGATTCAGCAATTAATGATGATGCTA 2657

DE 657 CTTTGGTGGATTCAGGATTCAGCAATTAATGATGATGCTA 714

QY 2657 CTTTGGTGGATTCAGGATTCAGCAATTAATGATGATGCTA 2717

DE 715 CTTTGGTGGATTCAGGATTCAGCAATTAATGATGATGCTA 773

Search completed January 18, 2003, 04:09:46
Job time: 2962.99 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

EM protein nucleic search, using frame_plus_p2m model
Run on: January 17, 2003, 18:27:24 : Search time 1840.65 seconds
(without alignments)
7320.587 Million cell updates/sec

Filter: US 0.9 0.26 459A 45
Perfect score: 4278
Sequence: 1 MWLGICFIATIAVILLESFT.....TMDKQKWNUSMETSNEEK 832

Scoring table:
Knap 10.0, Xgapext 0.5
Ygap 10.0, Ygapext 0.5
Egap 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searches: 16134066 seqs, 809774376 residues
Total number of hits satisfied chosen parameters: 42369342

Minimum db seq length: 0
Maximum db seq length: 200000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
Model: frame_plus_p2m model -DEV=win
C:\zoo\2.1\uspro\sp\seq\us09026459a\runat_16012003_182724\hpf_query.fasta_11447
-lib=EST -qfmt=lastup -suffix=est -minmatch=0.1 -looptcl=0 -looptext=0
-units=bits -start=1 -end=1 -matpfx=pl0sm62 -trans=human40.cdi -list=45
-localign=200 -thr_score=pt -thr_max=100 -thr_min=0 -align=15 -mode=local
-outfmt=pt -norm=exc -heapsize=500 -minlen=0 -maxlen=200000000
USER=us09026459a -seqn=1 -14595 -trans=16012003_182724_23322 -ncpu=6 -cpu=3
No XCOPY -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
WARN_TIMEOUT=40 -THREAPUS=1 -XAPCP=10 -XGAPEXT=0.5 -EGAPB=6 -EGAPEXT=7
YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database: EST
1: em_est.hum.*
2: em_est.hum.*
3: em_est.hum.*
4: em_est.hum.*
5: em_est.hum.*
6: em_est.hum.*
7: em_est.hum.*
8: em_est.hum.*
9: qb_est1.*
10: qb_est1.*
11: qb_est1.*
12: qb_est1.*
13: qb_est1.*
14: qb_est1.*
15: em_est.hum.*
16: em_est.hum.*
17: qb_est1.*
18: em_est.hum.*
19: em_est.hum.*
20: em_est.hum.*
21: em_est.hum.*
22: em_est.hum.*
23: em_est.hum.*
24: em_est.hum.*
25: em_est.hum.*
26: em_est.hum.*
27: em_est.hum.*

Result No. Query Match Length DB ID
1 1551 36.3 1051 13 RM561084
2 1473 34.4 1364 11 AK011246
3 1448 33.8 881 14 R0220275
4 1339 31.3 1001 13 RM473603
5 1254 29.3 727 9 AL597811
6 1236.5 28.9 880 14 R0222227
7 1208 28.2 875 13 RM453724
8 1111 26.0 1023 13 RM450041
9 1056.5 24.7 914 9 AA763411
10 1043 24.4 792 9 AL563757
11 1037 24.2 605 9 AL599633
12 1030 24.1 694 10 R0202846
13 1016 23.7 623 10 RM584181
14 1013 23.7 871 9 AA763485
15 1004 23.5 922 12 R0252543
16 978 22.9 588 10 RM468234
17 962 22.5 745 13 R1151568
18 959 22.4 594 10 AV715533
19 955 22.3 584 12 R0144857
20 947.5 22.1 734 12 R0144857
21 935 21.9 539 9 AL592799
22 908 21.2 757 10 R0539278
23 882 20.6 572 9 AL129941
24 879.5 20.6 734 13 R1522582
25 869.5 20.3 625 12 R0244088
26 868 20.1 597 12 R033246
27 844 19.7 484 10 R0168095
28 833.5 19.5 646 13 RM490416
29 824 19.3 493 19 AW502887
30 815.5 19.1 522 9 AA072789
31 781 18.3 2083 13 RM456728
32 751 17.6 453 10 AW503514
33 727 17.0 832 9 AL646038
34 683 16.0 438 12 R027047
35 681.5 15.9 759 14 R02304
36 667 15.6 651 13 R1696841
37 659 15.4 380 10 AW51604
38 658 15.4 462 9 AL606529
39 633 14.8 425 10 R0285796
40 628 14.7 672 13 R1868459
41 626 14.6 521 12 R0466536
42 625.5 14.6 387 12 R0998490
43 613 14.3 415 9 AL745644
44 611 14.3 410 12 R0560103
45 609 14.2 411 9 AL745643

RESULT 1
RM561084
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

RM561084
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
1051 bp mRNA Homo sapiens cDNA clone IMAGE550905
5' mRNA sequence
RM561084
RM561084.1 G1:18806043
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1051)
NIB-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

msAspArgSerAlaArgIleCatalysGlnIleGluAsnAspThrArgIleCysGluVal 180


```

5' mRNA sequence.
BM450031
BM450031.1 GI:18499071
EST.
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 1023)
NIH-MGC http://mgc.nhl.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Published (1996)
Contact: Robert Strausberg, Ph.D.
Email: eqa@nsl.nhl.nih.gov
Tissue Procurement: ATCC/DCID/DTP
cDNA library preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LUMC)
DNA Sequenced by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LAM2294 row: 4 column: 22
High quality sequence stop: 643.
FEATURES
Source
1..1023
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/label="IMAGE:5528037"
/clone_lib="NIH-MGC-72"
/tissue_type="melanotic melanoma"
/lab_host="pBIOH (phage-resistant)"
/notes="organ skin, Vector: pCMV SPORT6; Site 1: NotI;
Site 2: SalI. Cloned unidirectionally. Primer: Oligo dT
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 144 a 211 c 197 g 281 t
ORIGIN
Altimet Scores: 5,41e-109 Length: 1023
Prod. No.: 1111-00 Matches: 216
Score: 1111-00
Percent Similarity: 100.00% Conservations: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.97% Indels: 0
GB: 13 Gaps: 0
us-09-026-459a-35 (1-842) x BM450031 (1-1023)
67 LysValLysAsnIleAspSerValPheGlyIleValThrAlaTyrLysAspLePro 636
|||||
68 2 AAGAGCAAGAAATACAGCTTAAATTCAAAATCATTTGAACAGCATACAGGATCTTCT 61
|||||
69 67 AATATATATATATATATATATATATATATATATATATATATATATATATATAT 656
|||||
70 62 CATGCTGTTTCAAGAGATATATATATATATATATATATATATATATATATATAT 121
|||||
71 657 TTTValPheTyrAsnSerValPheMetGlnArgGlyCysThrAsnIleCysGlnTyrAla 676
|||||
72 122 ATAGTATATATATATATATATATATATATATATATATATATATATATATATATAT 181
|||||
73 677 SerThrArgProThrLeuSerProTleProHisIlePheArgSerProTyrLysPhe 696
|||||
74 182 AACACAGACCCGCTACCTTGTCATCAAAACCTTCACTTCCCTTCCAGGCTTCAAGTTT 241
|||||
75 677 ProSerSerProLeuArgTleProGlyCysAspTleTyrSerProLeuLysSerPro 716
|||||
76 242 CTTAGTTACAGCTTACGATCTGAGAGAGAAATCTATATATATATATATATATATAT 301
|||||
77 717 TTTGlySerSerGlyLeuProThrProThrProThrLysMetThrProTyrSerArgIleLeu 736
|||||
78 302 TATATAAAATTCACAGAGCTTGTAAACCAACCAACCAACCAACCAACCAACCAACCA 361
|||||
79 737 ValSerIleGlySerPheGlyThrSerGluLysPheGlnLysIleAsnGlnMetVal 756
|||||

```

```

|||||
362 GTATCAATGGTCAATCATTCGGGACATTCAGAAATTCAGAAATTCAGAAATTCAGAAAT 421
|||||
757 CysAsnSerAspArgValIleCulysArgSerAlaCileGlySerAspProProLysProLeu 776
|||||
76 422 TGTAAATAGAGAGATATATATATATATATATATATATATATATATATATATATATAT 481
|||||
777 TGTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 756
|||||
78 482 AAAAAATACGCTTGTATATATATATATATATATATATATATATATATATATATATATAT 541
|||||
79 542 GATAGATCAATATTCATAGAGAAATTCAGAAATTCAGAAATTCAGAAATTCAGAAAT 601
|||||
817 TysGlnLysMetAspSerMetAspThrSerAspGlyGlyGlyGlyGlyGlyGlyGlyGly 832
|||||
82 602 AATGAGAAAAATATATATATATATATATATATATATATATATATATATATATATATATAT 649
|||||
RESULT 9
AA763411/c
AA763411/c
DEFINITION
Mus musculus
house mouse.
ACCESSION
AA763411
VERSION
AA763411.1
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 914)
REFERENCE
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, L.,
Geisler, S., Kucuba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Willson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
COMMENT
Contact: Maria M. Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available for public use through LLNL; contact the
IMAGE Consortium (image.llnl.gov) for further information.
MGI:661286
Seq primer: 28m13 rev2 ET from Amersham
High quality sequence stop: 422.
Location/Qualifiers
1..914
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGE:1247598"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/notes="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Honaido."
BASE COUNT 257 a 178 c 217 g 261 t
ORIGIN

```



```

QY 244 AspLysSerLeuValThrAspSerLeuAspSerPheGlnArgThrProValArgValMet 263
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 451 GATAAAGTCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 432
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 264 SerAsnLeuAspGluValAsnValLeuProProHisThrProValArgThrValMet 283
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 431 AATAACCTTCCGACACACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 284 AsnThrIleGlnGlnMetMetIleLeuAsnSerAlaSerAspGlnProSerSerAsn 303
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 471 GATATTTCCACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 304 LeuIleSerIlePheAsnAsnGlnSerValAsnProGlySerGlnSerIleGlnGlySerVal 323
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 411 YTGTTTCTTTTATGACGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 324 LysAspIleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 343
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 451 AAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 192
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 344 GluIleGlnSerGlnArgThrLeuGlyValArgValMetGlnGlnGlnGlnGlnGlnGln 363
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 191 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 363 MetLeuLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 382
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 131 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 383 AspAsnIlePheIleSerLeuLeuAlaGlnValValValValValValValValValVal 402
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 71 AAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 403 SerArgSer 405
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 11 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 11
LOCUS AL599633
DEFINITION DKEZp13p092_1 313 (313bp, m, hloc2) B-mos sapiens cDNA clone
VERSION AL599633
KEYWORDS AL599633.1 GI:15162921
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
Koehler, K., Beyer, A., Mewes, H.W., Well, B. and Wiemann, S.
EST (Koehler, K., Beyer, A., Mewes, H.W., Well, B. and Wiemann, S.)
Unpublished (1999)
Contact: Koehler K
MIPS
Am Klipferspitze 19a D-92152 Martinsried, Germany
This is the 5' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DKEZp13p092) is available at the PCR in Berlin
Please contact the 5726 Resource Center, Heidenhain 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzp1.de
Location/Qualifiers
1. 605
/organism="Homo sapiens"
/lab_host="35-926"
/lab_project="459a"
/lab_stage="adult"
/lab_type="BHT08"

```

FEATURES
SOURCE

```

/Note=Vector: pStripEx2; Site_1: SfiI; Site_2: SfiI;
cDNA-collection: 428 3 92 c 105 4 179 1
BASE COUNT
ORIGIN
Alignment Scores:
  233c 101      length: 605
  1037 00      matches: 201
  100.00%     conservation: 0
  100.00%     mismatches: 0
  24.24%     indels: 6
  9          gaps: 0
DB:
US-004-026-459a-35 (1-832) X AL599633 (1-605)
QY 179 GluValLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 198
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 1  CAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 60
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 199 AsnPheIlePhePheMetAsnSerLeuGlyLeuValThrSerAsnGlyIleGlnGlnGlnVal 218
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 61  AATTTTATACCTTTTATGAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 219 GluAsnLeuSerLysArgThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 238
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 121 GAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 239 LeuPheLeuAsnGlnAspLysThrLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 258
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 181 TTATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 259 ArgPheProArgLysSerAsnLeuAspGlnGlnValAsnValIleProProHisThrPro 278
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 241 AGAATCAACAGCAAGAAAGTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 279 ValArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAsp 298
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 301 GTTAGGACCTGTTGATGACACATATCCCAATTAATGATGATGATGATGATGATGATGATGATGAT 360
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 299 GlnProSerGlnAsnLeuIleSerIleThrAsnAsnGlnGlnGlnGlnGlnGlnGlnGln 318
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 361 CAACCTTCAGAAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 319 IleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 338
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 421 ATACTCAAAAGAGTCAAGAGATATAGGATATATCTTTTAAAGAGAAATTTGTTAAAGAGTGTG 480
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 339 GlyGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 358
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 481 GAAATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGAT 540
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 359 ArgValMetGlnSerMetLeuLysSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 378
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 541 CCATTAATGCAATTCATCTCTTAAATCAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 379 Lys 379
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1b 601 AAA 603
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
RESULT 12
LOCUS BE082846/c
DEFINITION BE082846_1 604 (604bp, m, hloc2) Homo sapiens cDNA, 798A sequence.
VERSION BE082846
KEYWORDS BE082846.1 GI:8473151
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
Dias Neto, E., Garcia-Cortea, R., Velazquez, Almeida, S., Arbones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zagó, M.A., Bordin, S., Costa, P.F.,

```


1. The first part of the document is a letter from the President of the United States to the Congress, dated January 1, 1861. It is a very important document, as it contains the President's message to the Congress at the beginning of his first term. The letter is written in a formal, dignified style, and it is one of the most important documents in American history.


```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528037"
/clone_lib="NH_MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/vector="pMD19 (+) SK(-) Vector"
/site_2="SalI: Cloned unidirectionally. Primer: oligo dr.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT      334 a 211 c 197 g 281 t
ORIGIN
Query Match      29.2% Score 910.2; DB 13; Length 1023;
Best Local Similarity 98.4%; Pred. No. 4.9e-165;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

QY 1701 CAAGATGACATATAGAGCTTAAATTTAAATTTATTTGATACAGCATACAGAGATCTTCC 1760
DB 1 CAAGATGACATATAGAGCTTAAATTTAAATTTATTTGATACAGCATACAGAGATCTTCC 60
QY 1761 TCAAGTGTGTCAGATACATATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 1820
DB 61 TCAAGTGTGTCAGATACATATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 120
QY 1821 TATAATATTTATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1880
DB 121 TATAATATTTATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 1881 TTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1940
DB 181 TTGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 1941 TCTTATTTTCACTTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 2000
DB 241 TCTTATTTTCACTTATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 300
QY 2001 ATATATAATTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2060
DB 301 ATATATAATTTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 2061 AGTATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 2120
DB 401 AGTATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 420
QY 2121 AGTATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 2180
DB 421 AGTATCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 480
QY 2181 GAAATTAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 2240
DB 481 GAAATTAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 540
QY 2241 AGATATATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 2300
DB 541 AGATATATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 600
QY 2301 AAAGATAGAAAATGATATATATATATATATATATATATATATATATATATATATATAT 2360
DB 601 AAAGATAGAAAATGATATATATATATATATATATATATATATATATATATATATATAT 660
QY 2361 TAAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 2420
DB 661 TAAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 720
QY 2421 TAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 2480
DB 721 TAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 780
QY 2481 AAAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 2540
DB 781 AAAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCA 840

```

RM561083 1051 bp. mPRA library EST 20 FEB 2002
 AGENCOURT 6566133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550905
 5' mRNA sequence.
 RM561083
 BMS61083.1 GI:18806043
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (Bases 1 to 1051)
 NIH-MGC http://mml.nhlbi.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bms.fda.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 cDNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution. Information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.llnl.gov
 Plate: LINL2263 row: m column: 13
 High quality sequence stop: 737.
 FEATURES
 location/Qualifiers
 1..1051
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5550905"
 /clone_lib="NH_MGC_67"
 /tissue_type="retinoblastoma"
 /note="Ordan: eye; Vector: pCMV-Sport6; Site 1: Not;
 Site 2: SalI. Cloned unidirectionally. Primer: oligo dr.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."
 BASE COUNT 330 a 221 c 197 g 299 t 4 others
 ORIGIN
 Query Match 29.2% Score 909.4; DB 13; Length 1051;
 Best Local Similarity 98.4%; Pred. No. 4.9e-165;
 Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;


```

143 121 AATTTTAAATTTTCTGAAATACAAATATTTTCAATGCTTTTATTCGGCTGGCTCTT 180
QY 1049 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1098
143 143 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 240
QY 1099 TGGTCTTTCGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 1158
143 241 TGGTCTTTCGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 300
QY 1159 ATGCAAAATTTTATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAAT 1218
143 401 ATGCAAAATTTTATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAAT 360
QY 1219 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1278
143 461 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 420
QY 1279 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1338
143 421 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 480
QY 1339 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1398
143 481 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 540
QY 1399 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1458
143 541 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 600
QY 1459 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1518
143 601 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 660
QY 1519 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1578
143 661 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 720
QY 1579 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1638
143 721 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 780
QY 1639 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1698
143 781 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 840
QY 1699 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1757
143 841 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 900
QY 1759 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 1809
143 901 GAGGTCTTAATGGGCAATATATGCAATATGATCTCAGAAATCTTGATCTGGAACAGAT 953

RESULTS 4
LOCUS B04.49221 1081 bp mRNA linear EST 24 MAY 2002
DEFINITION AGC050827_776552 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015480
5' mRNA sequence.
ACCESSION B04.49221
VERSION B04.49221.3 G1:2178297
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1081)
TITLE NIH_MGC ht tp://imgc.ncbi.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: egapbs@mail.nih.gov

```

Tissue Procurement: AICC
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://imgc.llnl.gov
Plate: LLNL3211 row: k column: 01
High quality sequence stop: 742.

FEATURES

source

Location/Qualifiers

1..1081

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:6015480"

/clone_lib="NIH_MGC_92"

/issue_type="embryonal carcinoma, cell line"

/lab_host="DH10B (phage-resistant)"

/note="organ: testis; vector: pCMV-Sp-R1; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally; origin of primed;

Average insert size: 2.5 kb. Library enriched for

full length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

BASE COUNT 437 a 190 c 183 g 471 t

ORIGIN

Query Match 27.5% Score 856; Db 14; Length 1081;

Best Local Similarity 99.1% Pred. No. 9, w 100;

Matches 892; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

QY 2215 GAAGCAGATGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2274

Db 1 GAGGACAGTGAAGTAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 60

QY 2275 ATGACTTCTACTCGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2334

Db 61 ATGACTTCTACTCGAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 120

QY 2335 AACAGGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2394

Db 121 AACAGGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 180

QY 2395 CATGTCTCTCAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2454

Db 181 CATGTCTCTCAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 240

QY 2455 ATATCTTCAGTCTTTTGTGATATGATGATGATGATGATGATGATGATGATGATG 2514

Db 241 ATATCTTCAGTCTTTTGTGATATGATGATGATGATGATGATGATGATGATGATG 300

QY 2515 AAGCCATTTCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2574

Db 301 AAGCCATTTCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 360

QY 2575 TGGCATTTAAAAAGTCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2634

Db 361 TGGCATTTAAAAAGTCTGACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 2635 TGGATAGTAAATGAGGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2694

Db 421 TGGATAGTAAATGAGGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 480

QY 2695 TGGTCTTTCGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 2754

Db 481 TGGTCTTTCGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 540

QY 2755 TTTTTHAATTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2814

Db 541 TTTTTHAATTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 600

QY 2815 TTTTTHAATTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 2874

Db 601 TTTTTHAATTTAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 660

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MCC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LLAM12426 row: 1 column: 04

High quality sequence start: 17

High quality sequence stop: 734.

FEATURES

Location/Qualifiers

1..1091

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NIH_MCC_92"

/issue_type="embryonal carcinoma, cell line"

/lab_host="pH10B (phage-resistant)"

/note="organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; cloned unidirectionally; oligo-dT primed.

Average insert size 2.5 kb. Library enriched for

full-length clones and constructed by Life Technologies

Note: this is a NIH_MCC Library."

449 a 178 c 288 t

BASE COUNT

Query Match 24.2% Score 753.6; DB 13; Length 1001;

Best Local Similarity 95.0% Pred. No. 4.8e-135;

Matches 800; Conservative 0; Mismatches 39; Indels 3; Gaps 2;

QY 7 ATGTCAGAGCTGTTGAAGAGTATATGTTGTTGACCTTCAGCAATTC-GGAAGA 65

DB 77 AAGCAAGAGCTGTCGAAGAGTATGATGTTGCTTCAGCTTCAGCAAAATCGCAAG 136

QY 66 GATATGCGAACTAT 125

DB 137 GACAGCTGAGCTAT 196

QY 126 TGCATGCTGCTAAAGCTTCCTGCGACACATTTTATAGCTTAAGCGGAGTATTACA 185

DB 197 TGTATGCTGCTAAAGCTTCCTGCGACACATTTTATAGCTTAAGCGGAGTATTACA 256

QY 186 AATGAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 245

DB 257 AATGAGAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 316

QY 246 TAAACTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCT 305

DB 417 TAAACTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCT 376

QY 406 TGTTCACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCT 365

DB 477 TGTTCACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCTGCTACCT 436

QY 466 AGAAT 425

DB 437 AGAAT 496

QY 426 GGTGCAAAAGCTAT 485

DB 497 GGTGCAAAAGCTAT 556

QY 486 AATTAAT 545

DB 557 AATTAAT 616

QY 546 AATTAAT 605

DB 617 AATTAAT 676

QY 606 TAT 665

DB 677 TAT 736

QY 666 TGTAT 725

DB 736 TGTAT 796

DB 737 TGTAT 796

QY 726 GATTTTAAATTCAGCAAGTAT 783

DB 797 GATTTTAAATTCAGCAAGTAT 856

QY 784 TGCACAGTCAATTCAGCAAGTAT 843

DB 857 TGCACAGTCAATTCAGCAAGTAT 916

QY 844 AA 845

DB 917 AA 918

RESULT 6

BM545881

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1..1122

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NIH_MCC_125"

/lab_host="pH10B"

/note="organ: ovary (pool of 3); Vector: pCMV-SPORT6;

Site_1: EcoRV (destroyed); Site_2: NotI; RNA source pool

of three ovaries, from females ranging in age from 48 to

49 yo. Library is oligo-dT primed and directionally cloned

(RcoRV site is destroyed upon cloning). Average insert

size 2.1 kb. Insert size range 1-4.5 kb. Library is

normalized and enriched for full-length clones and was

constructed by C. Gruber (Invitrogen). Research Genetics

tracking code 046."

BASE COUNT 360 a 185 c 216 g 359 t 2 others

ORIGIN

Query Match 24.2% Score 752.4; DB 13; Length 1122;

Best Local Similarity 95.6% Pred. No. 8.2e-145;

Matches 827; Conservative 0; Mismatches 42; Indels 6; Gaps 9;

QY 2080 TTCCGCACTTCTCAGAGTTCAGCAAGTATATATATATATATATATATATATATATAT 2139

DB 240 TTGTAGACTTCTCAGAGTTCAGCAAGTATATATATATATATATATATATATATATAT 299

QY 2140 CTCAAAAGAGTCTGAAGAGCAAGCTCTCTAAACCACTGAAAAAATATAGTTTAT 2199

DB 300 CTCAAAAGAGTCTGAAGAGCAAGCTCTCTAAACCACTGAAAAAATATAGTTTAT 359

/lab_host="DH10B (T1 phage-resistant)"
 Note: Organism: primate; Vector: pDNR-LTR (Clontech);
 Site 1: SII (qccgcctggcc); Site 2: SII (qccattatggcc
); Double-stranded cDNA was prepared from cell line RNA.
 5' and 3' adaptors were used in cloning as follows: 5'
 adaptor sequence: 5'-CACGCCCATTAAGCC-3' and 3' adaptor
 sequence: 5'-ATTTCACAGGAGGAGGACAGG-3' (40)NN-3'
 (where N = A, C, G, or T). Average
 insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
 contained inserts by PCR. This library was enriched for
 full-length clones and was constructed by Clontech
 laboratories (Palo Alto, CA). Note: this is a NIH-MGC
 library."

BASE COUNT 252 a 148 c 112 g 271 t
 ORIGIN
 Query Match 22 58; Score 699.6; DB 12; Length 823;
 Best Local Similarity 94.8%; Pred. No. 1,26-124;
 Matches 779; Conservative 0; Mismatches 34; Indels 9; Caps 5;
 QY 2241 AGGAGAGTCAAAATTTCAAGAAATGGAGAAATTTCTATTTGAAATGAATGAA 2400
 DB 1 AGGAGAGTCAAAATTTCAAGAAATGGAGAAATTTCTATTTGAAATGAATGAA 60
 QY 2401 AAGACACAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
 DB 61 AAGACACAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 2461 GAGATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAA 2420
 DB 121 GAGATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAA 180
 QY 2421 TAACCTTCTGATGATTTTATGATGATTTTATGATTTTATGATTTTATGATTTTAT 2480
 DB 181 TAACCTTCTGATTTTATGATTTTATGATTTTATGATTTTATGATTTTATGATTTTAT 240
 QY 2481 AAAAGTCCAGATGCAATTTCTGCTGATTTCTGCTGATTTCTGCTGATTTCTGCTGAT 2540
 DB 241 AAAAGTCCAGATGCAATTTCTGCTGATTTCTGCTGATTTCTGCTGATTTCTGCTGAT 300
 QY 2541 TATTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2600
 DB 301 TATTATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
 QY 2601 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2660
 DB 361 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 2661 GAGATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAAATTT 2720
 DB 421 GAGATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAAATTTGATGAAATTT 480
 QY 2721 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2780
 DB 481 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 QY 2781 TATCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2840
 DB 541 TATCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 2841 TATCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2897
 DB 601 TATCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 QY 2898 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2957
 DB 661 ATTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 717
 QY 2958 AAAATTTTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3017
 DB 718 AAAATTTTCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 776
 QY 3018 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3059

DB 777 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 816
 RESULT 10
 AL563757/c
 LOCUS
 DEFINITION
 AL563757 Homo sapiens cDNA clone Q500007YA05.3
 prime, mpna sequence.
 ACCESSION
 AL563757
 VERSION
 AL563757.1 GI:12913464
 KEYWORDS
 EST.
 SOURCE
 human.
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 792)
 Li, W.B., Gruber, C., Jessel, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 JOURNAL
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequençage
 BP 191 91005 Evry cedex - France
 Email: seqref@genoscope.cns.fr, web: www.genoscope.cns.fr.
 Location/Qualifiers
 1. 792
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="Q500007YA05"
 /clone_lib="LTL_NF1001_NHC4"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organism: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-cligod(T) primer. Five prime end
 enriched, double-stranded cDNA was digested with NotI and
 cloned into the NotI and EcoRV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed
 by Life Technologies. Contact: Feng Liang Life
 Technologies, a division of Invitrogen 9800 Medical Center
 Drive, Rockville, Maryland 20850, USA Fax: (1) 301 610
 8371 Email: fliang@lifetech.com url:
 http://fulllength.invitrogen.com"
 BASE COUNT 246 a 148 c 117 g 255 t 26 others
 ORIGIN
 Query Match 22 48; Score 498.2; DB 9; Length 792;
 Best Local Similarity 93.6%; Pred. No. 2,3e 124;
 Matches 722; Conservative 23; Mismatches 23; Indels 3; Caps 2;
 QY 282 TAAATAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 741
 DB 792 TAAATAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 733
 QY 342 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 401
 DB 732 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
 QY 402 AATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 461
 DB 672 AATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
 QY 462 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 521
 DB 612 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 553
 QY 522 AGCATACCAAGAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 581
 DB 552 AGCATACCAAGAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 641
 QY 582 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 641
 DB 492 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 433

100

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM nucleotide nucleotide search, using SW model

Run on: January 16, 2003, 15:20:22 : Search time 2966.48 seconds
(without alignments)
184669.506 Million cell updates/sec

Library: US-09-026-459A-48

Perfect score: 4383

Sequence: 1 GCGGCAAGCGGCGGAAAAA.....AAACAGCAATATTCATAGT 3383

Scoring table:

IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searches: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 4 (top 4)

Minimum DB seq length: 0

Maximum Match 100%

Listed first 4% summaries

Post processing: Minimum Match 0%

Maximum Match 100%

Listed first 4% summaries

Database:

EST: *

1: cm_estb1*

2: cm_estb2*

3: cm_estb3*

4: cm_estb4*

5: cm_estb5*

6: cm_estb6*

7: cm_estb7*

8: cm_estb8*

9: cm_estb9*

10: cm_estb10*

11: cm_estb11*

12: cm_estb12*

13: cm_estb13*

14: cm_estb14*

15: cm_estb15*

16: cm_estb16*

17: cm_estb17*

18: cm_estb18*

19: cm_estb19*

20: cm_estb20*

21: cm_estb21*

22: cm_estb22*

23: cm_estb23*

24: cm_estb24*

25: cm_estb25*

26: cm_estb26*

27: cm_estb27*

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.4	26.9	1023	13	BM4500331 AGENCOURT
2	909.4	26.9	1051	13	BM561083 AGENCOURT
3	896	25.3	1061	14	B0439221 AGENCOURT
4	762.4	22.2	1122	13	BM543681 AGENCOURT
5	725.4	21.4	727	9	AL597811 DKFP3130
6	699.6	20.7	824	12	B0510551

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

7	692.2	20.5	767	12	B0529466
8	691.8	20.4	972	13	BM466194
9	678	20.0	845	13	BM454724 AGENCOURT
10	660	19.5	672	13	BT668559
11	643.6	19.0	797	10	BE539278
12	642.2	19.0	621	12	B0516219
13	638.8	18.9	1364	11	AK011246
14	617	18.2	618	13	BM264009
15	614	18.1	694	10	BE082846
16	596.4	17.6	922	12	B3254343
17	580	17.1	580	9	AL599105
18	575.2	17.0	914	9	AA763411
19	566.4	16.7	588	10	AA368234
20	564.9	16.7	881	14	B0223275
21	558	16.5	871	9	AA763485
22	530.2	15.7	598	10	AA582967
23	530	15.7	539	9	AL692790
24	528	15.6	572	9	AL120911
25	528	15.6	605	9	AL599633
26	524	15.5	623	10	AA584181
27	521.2	15.4	792	9	AL563757
28	518	15.3	832	9	AL646048
29	517	15.3	1154	14	BM802736
30	513.2	15.2	594	10	AV715534
31	512.4	15.1	709	9	AL646097
32	512.2	15.1	734	12	BF144857
33	505.2	14.9	736	13	BT151568
34	493.8	14.6	525	9	AA582955
35	493.4	14.6	597	12	BE931236
36	486.4	14.4	584	12	B0149050
37	484.4	14.3	494	10	AA502887
38	481.4	14.2	484	10	BE168095
39	480.2	14.2	659	13	BT821116
40	478.6	14.1	615	9	AA959846
41	474.4	14.0	734	13	B1525982
42	459.2	13.6	497	12	BE931236
43	457	13.5	457	9	AL093215
44	457	13.5	732	9	AL598766
45	455.4	13.5	643	13	BT554937

ALIGNMENTS

RESULT 1
BM450041
LOCUS: BM450031 1023 bp mRNA Library EST 05 FEB 2002
DEFINITION: ACEN0001694544 NIH_MGC_74 Homo sapiens cDNA clone IMAGE:5549637
5% mRNA sequence.

ACCESSION: BM450031

VERSION: BM450031.1 GI:184996071

KEYWORDS: EST.

SOURCE: Human.

ORGANISM: Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 1023)

NIH-MGC http://mgi.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: rgs@pds.rci.nih.gov

Tissue Procurement: ATCC/ATCC/ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1238 row d column 22

High quality sequence stop: 64%

Location/Qualifiers

1..1023

and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCGATTATGCT-3' and 3' adaptor sequence: 5'-ATTCTAGACCGCCGCGGACATG-dT(30)NN-3' (where N = A, C or G and NN = A, C, G, or T). Average insert size 1.76 kb (range 0.9-4.6 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC library."

BASE COUNT	210 a	108 c	104 g	259 t	
CHISIN					
Query Match	19.0%	Score 642.2	DB 12	Length 681	
Best Local Similarity	68.4%	Prod No 2 1e-118			
Mismatches	670	Conservative	8	Mismatches	3
				Indels	3
				Gaps	2
QY	2697	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	2764		
DB	1	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	60		
QY	2757	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	2816		
DB	61	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	110		
QY	2817	TATAGATTTTATTTTATTTATTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	2876		
DB	120	TATAGATTTTATTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT	179		
QY	2877	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	2936		
DB	180	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	239		
QY	2937	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	2996		
DB	240	TCGCAGCTCTGTTTATGCGGACATTAATATATCTTCACGCTCTTTTGTGGCGGATATAAATG	299		
QY	2997	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3056		
DB	300	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	359		
QY	3057	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3116		
DB	400	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	419		
QY	3117	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3176		
DB	420	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	479		
QY	3177	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3236		
DB	480	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	539		
QY	3237	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3294		
DB	540	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	599		
QY	3295	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3354		
DB	600	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	659		
QY	3355	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	3375		
DB	660	TCGCAGCTCTGTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT	680		
RESULT 14					
AF011246					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					

KEYWORDS

Source

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

HTC: CAP Trapper.
Mus musculus (strain: C57BL/6J) 10 days embryo cDNA library
clone: lib-100017422.
Mus musculus
Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 Carninci, P., and Hayashizaki, Y.
High efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (11), 1617-1630 (2000)
20499374
3 Carninci, P., Itoh, M., Aizawa, K., Nagata, S., Sasaki, H., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, P., Matsumoto, H., Sakauchi, S., Ikegami, T., Kashiwagi, K.,
Fujitawa, S., Inoue, K., Terajima, Y., Iwaki, M., Chijiwa, T., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Itoh, Y., Kira, A., and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system: a rapid
sequencing pipeline with 384 multiplexed capillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20540913
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aizawa, K., Iwaki, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, T., Hono, H., Kasukawa, I., Saito, K.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, J.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, H., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G., Tomita, M.,
Quackenbush, J., Schriml, L., Staudli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Haidar, P., Harsh, G., Blake, J., Hoffell, D., Bojunga, N.,
Carninci, P., de Bonaldo, M., Brownstein, M., Bult, C.,
Fleischer, C., Fujita, M., Caribol, M., Gustincich, S., Hill, D.,
Holmann, M., Hume, D., Kamiya, M., Lee, N., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., McKaerle, P., Nordone, P.,
Rind, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K., Suzuki, H.,
Toyooka, K., Wang, K., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21095660
5 (bases 1 to 1364)
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hara, A.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imoto, K., Ishii, Y., Itoh, M., Iwaki, M., Kasukawa, I.,
Kato, H., Kawai, J., Kojima, Y., Kono, H., Kondo, M., Koyama, S.,
Furubara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ocho, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, F., Sakai, K., Sakai, K., Sato, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogata, Y., Sasaki, H., Tagami, M., Taira, A., Takahashi, F.,
Tanaka, Y., Terajima, Y., Toya, T., Yamamura, I., Yamakawa, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and
Hayashizaki, Y.

from GRESTES PCR (U.S. Letters Patent application No. 196
7716 - Ludwig Institute for Cancer Research) profiles
into the p9C 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions.

BASE COUNT 411 A 129 C 127 G 226 T 1 others

Query Match 18.16; Score 614; DB 10; Length 694;

Best Local Similarity 97.48; Prod. No. 9.3e-113;

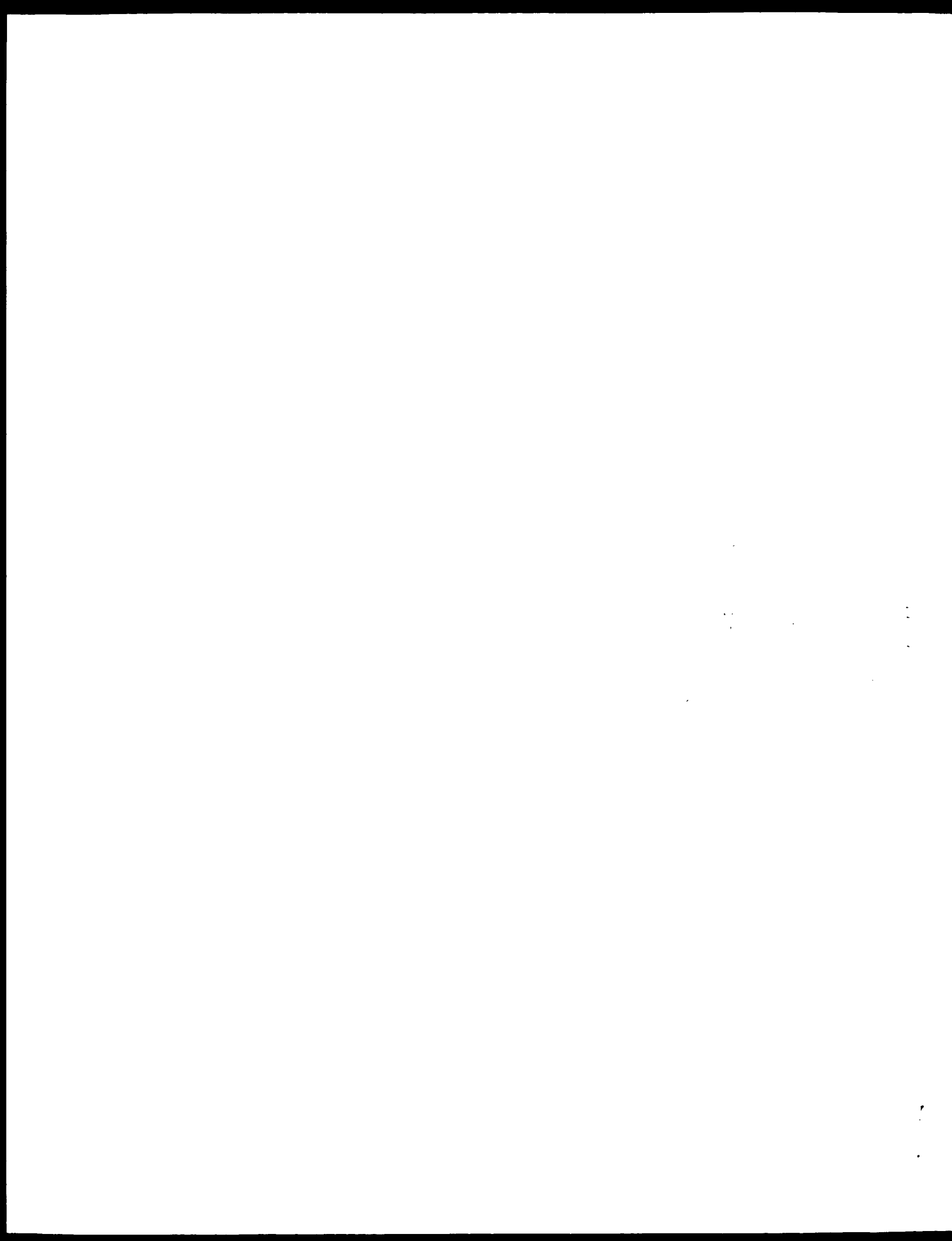
Matches 64%; Conservative 9; Mismatches 16; Indels 2; Gaps 2;

```

27 1044 TTTTAACTGCACTGATGATGCAAAAGAAATATATCTGAAAGAGGTGAAGGATATAGG 1103
    |||||
60 683 TATTTAAGCACTGACAGC AICCAAGAAAGATATAGTGAAGAGAGTGAAGGATATAGG 625
    |||||
27 1104 ATATATCTTAAAGAGAAATTTGCTAAAGTGTGAGGACACAGGTGTCGCAAAATGCAAC 1163
    |||||
60 624 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 565
    |||||
27 1164 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1223
    |||||
60 504 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 505
    |||||
27 1224 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1283
    |||||
60 504 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 445
    |||||
27 1284 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1343
    |||||
60 444 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 385
    |||||
27 1344 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1403
    |||||
60 484 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 425
    |||||
27 1404 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1463
    |||||
60 424 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 265
    |||||
27 1464 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1523
    |||||
60 264 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 205
    |||||
27 1524 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1583
    |||||
60 204 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 145
    |||||
27 1584 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1642
    |||||
60 144 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 85
    |||||
27 1644 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 1702
    |||||
60 84 ATATATCTTAAAGAGAAATTTGCTAGAGTGTGAGGACACAGTGTGTCGCAAAATGCAAC 25
    |||||
27 1704 AT 1705
    |||
60 24 AT 22

```

Search completed: January 16, 2003, 04:11:17
Job Time: 29.09.56.8008




```

1879 GAAVAAVTCATACATTCAGATCGCTGGCTCTTAATCTCTCCCTCCAGAAATATCAC 1938
QY 144 ThrAlaAlaAspMetTyrLeuSerProValArqSerProLysLysLysSerThrThr 563
1949 ATTCAGACAGATAATGATATCTTCTCCCTGAAGATCTCCAAAGAAAAGCTTCAACTAGC 1998
QY 164 ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLys 583
1999 GGTGTAAATTCATCTGTAATGTAGAGACACAAGCAACCTCAGCTTTCANATTCAGAG 2058
QY 184 ProLeuLysSerThrSerLeuSerLeuPheTyrLysValTyrArgLeuAlaTyrLeu 603
2059 CATTCAGAAATCAGCTCTCTTCTACCTCTTTATATAAAAGCTGATCGGCTAATCTATCT 2118
QY 604 ArgLeuAsnThrLeuLysGluArqLeuLeuSerGluLysProGluLeuLysIleIle 623
2119 GAGTAAATACATCTCTGAAAGCTCTCTGCTCAGCAGCAGCAGCAATAGACATATATC 2178
QY 624 TrpThrLeuPheGlnHisThrLeuGlnAsnGluTyrGluLeuMetArgAspArgHisLeu 643
2179 TGAAGCTTTTCCAGAGACATCTCTGAGATGATCAATTCATGAGACATAGAGATTTG 2238
QY 644 AspGlnIleMetMetCysSerMetTyrCysLysValLysAsnIleAspLeuLys 663
2249 GATTAATATATGATGCTGCTCAATGATGCGATATGCAATATGCAAGCTCAAGCTTAAA 2298
QY 664 PheLysIleIleValThrAlaTyrLysAspLeuProHisAlaValGlnGluThrPheLys 683
2299 TCAAAATCATCTCAACAGATACAAAGCACTCTCTCATGCTCTCAGAGACATTCAAA 2358
QY 684 ArgValLeuIleLysGluGluGluTyrAspSerIleIleValPheLysSerValPhe 703
2359 GGTCTTTTGATCAAAATAAAGATATATGATTTATATATATATATATATATATATATAT 2418
QY 704 MetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProThrLeuSer 723
2419 ATGACACACTGAAACAAATATTTCCAGTATGCTTCCACAGCGCCCTACCTTGCTCA 2478
QY 724 ProIleProHisIleProArqSerProTyrLysPheProSerSerProLeuArqIlePro 743
2479 GCAATATCTCACAATCTCTCAAGCGCTTCAAGCTTCTTCTAGTTCACCGCTACCGATCT 2538
QY 744 GlyCysAsnIleTyrIleSerProLeuLysSerProTyrLysIleSerGluGlyLeuPro 763
2539 GAGAGAAATCATATATTTTCCGCTGAAAGTCTCATATAATAATTTTCAGAGGCTCTGCCA 2598
QY 764 ThrProThrLysMetThrProArqSerArqIleLeuValSerIleGlyGluSerPheGly 783
2599 ATATTAATAAAAAGATCTGATAAGATCAACAAATCTTACTATCAATTCCTGCAATCCGG 2658
QY 784 ThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArqValLeuLys 803
2659 ATTTTCAGAGATTCATCAAAAAATATTCAGATGATATGATATGATATGATGCTGCTCAA 2718
QY 804 ArgSerAlaGluLysSerAspProProLysProLeuLysIleCysArqIleCysIleGlu 823
2719 AAGATGCTGAAAGAGCAAGCTCTTAACTACTGAAAGAACTAGCTTTGATATGAA 2778
QY 824 GlySerAspIleAlaAspGlySerTyrHisIleLeuProGlyCysPheGlnGlnLys 843
2779 GATATCAATCAAGATCAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 2838
QY 844 LeuAlaGluMetThrSerThrArqThrArqMetGlnLysGlnLysMetAsnAspSerMet 863
2839 CTGGTCAAAATCATCTCTTCTATCTCAACAGCAATTCAGAAAGCAAGATCAATCAATCA 2898
QY 864 AspThrSerAsnLysGlnLys 871
2899 GATATCTCAAAATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT

```

PESTLE 2
 AR098189

```

LOCUS AR098189 2994 bp DNA linear 14 FEB 2001
DEFINITION Sequence 3 from patent US 6074850.
ACCESSION AR098189
VERSION AR098189.1 GI:12807446
KEYWORDS Unknown.
SOURCE Unknown.
ORIGIN Unknown.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Antelmus, D., Gregory, R.J., and Willis, K.N.
TITLE Relinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 3 13-JUN 2000;
FEATURES Location/Qualifiers
source 1..2994
base_count 974 a 618 c 593 q 809 t
origin /organism="unknown"
Alignment Scores:
Pred. No.: 2,940-310 Length: 2994
Score: 4447.50 Matches: 871
Percent Similarity: 94.86% Conservativeness: 0
Best Local Similarity: 94.86% Mismatches: 0
Query Match: 98.92% Indels: 57
DB: 6 Gaps: 2
US-09-026-459A-49 (1-871) x AR098189 (1-2994)
QY 1 MetProProLysThrProArqLysThrAlaAlaThrAlaAlaAlaAlaAlaAlaPro 20
DB 139 ATGCGCGCCAAAACCCCGGAAAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 198
QY 21 ProAlaProProProProProProProProGluGluAspProGluGluAspSerGlyPro 40
DB 199 CGGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 258
QY 41 AspleuProLeuValArgLeuGluPheGluGluThrGluGluProAspPheThrAlaLeu 60
DB 259 GACCTGCTCTGTCAGGCTTGAGTTTGAGAGAAACAGAGAAACAGAGAAACAGAGAAAC 318
QY 61 CysGlnLysLeuLysIleProAspHisValArgGluArgAlaTrpLeuThrTrpLeuLys 80
DB 319 TGCACAGAAATTAAGATACACAGATCAATGACAGACAGAGCTTGGTAACTTGGAGAAA 378
QY 81 ValSerSerValAspLysValLeuGlyGlyTyrIleGlnLysLysLysLysLysLys 100
DB 379 GTTTCATCTGTGATGAGATGATGAGAGGTATATATCAAAAGAAAGAAAGAAAGTGGG 438
QY 101 TCCysIlePheIleAlaAlaValAspLeuAspGluMetSerPheThrPheThrGluLeu 120
DB 439 ATCTGTATCTTTATGTCAGCAGCTTGAGCTAGATGAGATGCTTATTCATTTATTCAGTA 498
QY 121 GlutLysAsnIleGluSerValHisLysPhePheAsnLeuLeuLysGlnIleAspThr 140
DB 499 CACAAAACATACAAATACAGATGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 558
QY 141 SerThrLysValAspAsnAlaMetSerArqLeuLeuLysLysLysLysValLeuPheAla 160
DB 559 AGTACCAATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 618
QY 161 LeuPheSerLysLeuGluArgThrCysGluLeuIleTyrLeuThrGlnProSerSer 180
DB 619 CTCTTTGAGTAAATGGAAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 678
QY 181 IleSerThrGluIleAsnSerAlaLeuValLeuLysValSerTrpIleThrPheLeuLeu 200
DB 679 ATATCTACTCAATTAATATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 738
QY 201 AlaLysGlyGluValLeuGlnMetGluAspAspLeuValIleSerPheGlnLeuMetLeu 220
DB 739 GCTAAAGGGAAGTATTACAAATGGAAGATGATGATGATGATGATGATGATGATGATGATGAT 798
QY 221 CysValLeuAspLysPheIleLysLysSerProProMetLeuLeuLysLysLysLys 240

```


[illegible]

RESULT 4	RD009728	RD009729	DNA	linear	147	JAN 2002
DEFINITION	tissue specific expression of retinoblastoma protein					

QY 181 IleSerThrGluIleAsnSerAlaIleuValIleuLysValSerTrpIleThrPheLeuLeu 200
 DB 544 ATAICTACTCAAAATAAATTGTGATTGGTGTCTAAAAGTTCTTGGAICACATTTTATTA 603
 QY 201 ALILysGlyGluValLeuGlnMetGluAspLeuValIleSerPheGlnLeuMetLeu 220
 DB 604 GCATAAGGCAAGTATTACAAATGGGAAGATGATCTGCTGATTTTCATTCAGTTAATGCTTA 663
 QY 221 CysValIleAspTyrPheIleLysLeuSerProProMetLeuLeuLysGlnProTyrLys 240
 DB 664 TGTGTCTTCAATATTTTATTAATATCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
 QY 241 Phe-----ClySer----- 243
 DB 724 ACACATTTTATATATTAATGCTTCACTGGAACACCTGAGGCAAGAGAGT 783
 QY 243 ----- 243
 DB 784 GCACGGCAACCAAAACACACAGAAATGATACAGAAATTTATTCAGATTCCTCTGCTAAAGAA 843
 QY 243 ----- 243
 DB 844 CATCAATTAATATATATGAGGTGAAATGTTTATTTCAAAAATTTTATACCTTTTAAAG 903
 QY 244 AsnSerLeuGlyLeuValIleThrSerAsnGlyLeuProGluValIleAsnLeuSerLysArg 263
 DB 904 AAATCTCTGGACTTGTAACTATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 963
 QY 264 TyrPheGlnIleTyrLeuLysAsnLysAspLeuAspAlaArgLeuPheLeuAspHisAsp 283
 DB 964 IACCAAGAAATTTATCTTTAAAAATAAAGATCTAGATGCAAGATATTTTTCGATCATGAT 1023
 QY 284 LysThrLeuGlnThrAspSerIleAspSerPheGluThrGlnArgThrProArgLysSer 303
 DB 1024 AAAATCTTCAGACIGATCTATAGACATTTTGAACACACAGAAACACCCAGCAAAAAGT 1083
 QY 404 AsnLeuAspGluGluValAsnValIleProProHisIleThrProValArgThrValMetAsn 323
 DB 1084 AAGCTTGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143
 QY 424 ThrIleGlnGlnLeuMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 343
 DB 1144 ACTACCAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
 QY 444 IleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIleLeuLysArgValLys 363
 DB 1204 ATTCTGATTTTAAACAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
 QY 464 AspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValIleClnClyCysValGlu 383
 DB 1264 GATATAGATATATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
 QY 294 IleLysSerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArgValMetGluSerMet 403
 DB 1324 ATGGATACACAGCAATCAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
 QY 404 LeuLysSerGluGluArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn 423
 DB 1484 CTAAATACAGAAAGCAAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1443
 QY 423 IlePheHisMetSerLeuLeuAlaCysAlaIleGluValIleMetAlaThrTyrSerArg 443
 DB 1444 ATTTTCTATATGCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
 QY 444 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProTrpIleLeuAsnVal 463
 DB 1504 AGTACATCTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1563
 QY 464 LeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPheIleLysAlaClnCly 483
 DB 1564 CTAAATTTAAAGCCCTTTATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1623
 QY 484 AsnLeuThrArgGlnMetIleLysHisIleGluArgCysGlnHisArgIleMetGluSer 503

DB 1624 AACTTCACAGCAAAATGATATAAAATTTAGAACGATGCGAACATCGAAATCAATCAATCC 1683
 QY 504 LeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAsnArgGln 523
 DB 1684 CMTGCAATGGCTCTCAGATTCACCTTTTATTTGATCTTATTTAAACAAATCAAAACACCGAGAA 1743
 QY 524 GlyProThrAspHisLeuGluSerAlaCysPheLeuAsnLeuProGluGlnAsnHis 543
 DB 1744 GAGCAAAATGATATATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 1803
 QY 544 ThrAlaAlaAspMetTyrLeuSerIleValArgSerIleLysLysLysLysLysLysLysLys 563
 DB 1804 ACTGCAATGATATGATTTCT 1863
 QY 564 ArgValAsnSerThrAlaAsnAlaGlnThrGlnAlaThrSerAlaPheGlnThrGlnLys 583
 DB 1864 CGTGAATTTCTAGTCAAAATGACAGACACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1923
 QY 584 ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrAlaGlnAlaTyrLeu 603
 DB 1924 CATTAAATTTATCTTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1983
 QY 604 ArgLeuAsnThrLeuGlnCysGluArgLeuLeuSerGlnHisPheGluLeuGlnHisIleIle 623
 DB 1984 CGCTAAATACACTTTTGTCAAGGCT 2043
 QY 624 TrpThrLeuPheGlnHisThrLeuGlnAsnGluTyrGluLeuMetAlaAspArgHisLeu 643
 DB 2044 TGGACCTTTTCTCAGACACACTTGTGAGAAATGATGATGATGATGATGATGATGATGATGATGAT 2103
 QY 644 AspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLysAsnIleAspLeuLys 663
 DB 2104 GACCAAAATTTATGATGTTTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2163
 QY 664 PheLysIleIleValIleThrAlaTyrLysAspLeuProHisAlaValGlnGluThrPheLys 683
 DB 2164 TTTCAAAATCTTGTAAACAGCATACAGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2223
 QY 684 ArgValLeuIleLysGluGluTyrAspSerIleIleValPheTyrAsnSerValPhe 703
 DB 2224 CGTCTTTTCTCAAAACAGCAGCATGATCTATTTATTTATTTATTTATTTATTTATTTATTT 2283
 QY 704 MetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProThrLeuSer 723
 DB 2284 ATGCAGAGATCTAAACAAATATTTTGTAGTATGTTCTCAACAGGCAATTTATTTATTTATTT 2343
 QY 724 ProIleProHisIleThrArgSerTyrLysLysIleProSerSerProIleClnArgIlePro 743
 DB 2344 CCAATACCTTCACATCTCTCTCAAGCCCTTCAAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2403
 QY 744 GlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIleSerGlnClyLeuPro 763
 DB 2404 GGAGGCAACATCTATTTTACCCCTTGAAGATGCTCAATATAAAATTCACAAAGCTCTCTCTCT 2463
 QY 764 ThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyGluSerPheCly 783
 DB 2464 ACACCAACAAAATGACTCCCAAGATCAAGATCTTTAGTATCAATTTGTTGATTTGATTTGAT 2523
 QY 784 ThrSerGluLysPheClnLysIleAsnGlnMetValLysAsnSerAspArgValIleLys 803
 DB 2524 ACTTCTGCAAGATCTCCAGAAAATAAAACAGATGCTCAATGCTGCTGCTGCTGCTGCTGCTGCT 2583
 QY 804 ArgSerAlaClnClySerAsnProProLysProLeuLysLysLeuArgPheAspIleGln 823
 DB 2584 AGAATGCTGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 2643
 QY 824 GlySerAspGlnAlaAspGlySerLysHisLeuProGlyGluSerLysPheGlnClnLys 843
 DB 2644 GATATGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2703
 QY 844 LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsnAspSerMet 863

[illegible][illegible]


```

1159 AAAACCTTCAGACGCTATTTCTATAGACGCTTTTAAACACAGAGAAACACGAAAAAGT 1218
1204 ASNLEUASPEGLUVALASNVALLLEPPTOPROHISLTHRPVALARQTHRVALLMETASN 1263
1219 AACCTTCAGACGCTGAAAGTAACTTCCTCACACACATCCAGCTTAGACCTGTTATGAAC 1278
1224 ThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 1343
1279 ACTATCCAACTTAATGATGATTTTAAATTCAGCAAGTATCAACCTTCAGAAAACTCG 1338
1344 IleSerIlyPheAsnAsnCysThrValAsnProIlySerGluSerIleLeuIlyArgValIly 1363
1349 ATTTCATATTTAACTATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1398
1364 AspIleCysIlyIlePheIlySerGluValAlaValAlaValAlaValAlaValAlaVal 1383
1399 GATAAGATATATCTTTAAACACAGAAATTCCTAAAGCTGCGGACACGGCTGCTCCAA 1458
1444 IleGlySerGlnAlaPheIlyLysLeuGlyValAlaValAlaValAlaValAlaValAla 1403
1459 ATTGGAATCAGAGGATACAAATTCGAGTTCGCTGTTATACCGAGTAATCGAATCCATG 1518
1504 LeuIlySerGluCysGlnAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 1523
1519 CTAAATTCAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1578
1624 IlePheHisMetSerLeuLeuAlaCysAlaValAlaValAlaValAlaValAlaValAla 1643
1679 ATTTCATATTTAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1638
1744 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProTrpIleLeuAsnVal 1763
1749 ASLACATCAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1798
1804 LeuAsnLeuIlyAlaPheAspPheIlyLysValIleGluSerPheIleLysAlaIleCly 1823
1809 CTAAATTTAAACCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
1844 AsnLeuThrArgGluMetIleLysHisLeuGluAlaCysGluHisArgIleMetGluSer 1903
1849 AACTTCAGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
1904 LeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu 1923
1949 CTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1878
1979 GAACTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1938
2044 ThrAlaAlaAspMetIlyLeuSerProValArgSerProIlyLysLysClySerThrThr 2063
2049 ACTGCAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1968
2064 ArgValAsnSerThrAlaAsnAlaThrAlaAlaThrSerAlaPheCysThrGlyCys 2083
2099 GAGGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2058
2144 ProLeuIlySerThrSerLeuSerLeuPheIlyLysValIlyArgGluAlaIlyLeu 2163
2149 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2118
2164 ArgLeuAsnThrLeuCysGlnAlaValAlaValAlaValAlaValAlaValAlaValAla 2183
2169 GAGGATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2178
2244 TrpThrLeuPheGlnIleThrLeuGlnAsnGluTrpGluLeuMetArgAspArgHisLeu 2263
2269 GAACTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2238
2344 AspGlnIleMetCysSerMetIlyGlyIleCysLysValIlyAsnIleAspLysLys 2363
2349 GAACTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2298

```

```

QY 664 PheCysIleIleValThrAlaIlyIlySerAspLeuProHisAlaValGlnGlnThrPheLys 683
DB 2299 TCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2458
QY 684 ArgValLeuIleLysGlnGlnGlnIlyIlyAspSerIleIleValPheIlyIlyAsnSerValPhe 703
DB 2359 CGTGTTCATCAAAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2418
QY 704 MetGlnArgLeuIlyThrAsnIleLeuGlnIlyAlaSerThrArgProProThrLeuSer 723
DB 2419 ATGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2478
QY 724 PheIlePheHisIlePheCysSerProIlyLysPhePheCysSerSerPheLeuArgAlaIlePro 743
DB 2479 GCAATACCTCACATGCTCCGAAAGCTTACAGTTCCTAGTTCACCTTACCGATTCCT 2538
QY 744 GlyGlyAsnIleIlyIleSerProLeuLysSerProIlyLysIleSerGlnGlyLeuPro 763
DB 2539 GAGGGAATATATATATTCACCTTAAAGATCCATATATATATATATATATATATATATATATAT 2598
QY 764 ThrProThrIlySerMetThrProAlaSerArgIleLeuValSerIleGlyGluSerPheCly 783
DB 2599 AATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2658
QY 784 ThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgValLeuLys 803
DB 2659 ACTTCGACAGCTCCGAAATTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2718
QY 804 ArgSerAlaGlnGlySerAspProProLysProLeuLysLysLysLeuArgPheAspIleGln 823
DB 2719 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2778
QY 824 GlySerAspGluAlaAspClySerLysHisLeuProGlyGluSerLysPheGlnGlnLys 843
DB 2779 GATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2838
QY 844 LeuAlaGlnMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsnAspSerMet 863
DB 2839 CTGGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2898
QY 864 AspThrSerAspLysGluGluLys 871
DB 2899 GATACCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2922
RESULT 11
LOCUS HUMRBS 4740 bp mRNA linear FRI 12 JUL 1995
DEFINITION Human retinoblastoma susceptibility mRNA, complete cds.
ACCESSION M13400
VERSION M13400.1 GI:190958
KEYWORDS retinoblastoma susceptibility.
SOURCE Homo sapiens (clone: RA-11.5) foetus retina cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Mammalia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 4740)
AUTHORS Lee, W.H., Rockstein, R., Hong, F., Young, L.J., Shaw, J.Y., and Lee, P.Y.
TITLE Human retinoblastoma susceptibility gene: cloning, identification,
and sequence
JOURNAL Science 235 (4794), 1394-1399 (1987)
MEDLINE 87149066
PUBMED 3823889
REFERENCE 2 (bases 1 to 480)
AUTHORS Lee, E.Y., Rockstein, F., Young, L.J., Lin, C.J., Rosenfeld, M.D., and
Lee, W.H.
TITLE Molecular mechanism of retinoblastoma gene inactivation in
retinoblastoma cell line Y79
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (16), 6017-6021 (1988)
MEDLINE 88320374
PUBMED 3413073
COMMENT Draft entry and computer-readable copy of sequence in [1] kindly
provided by E. Rockstein, 27-APR-1987.

```



```

QY 404 AsnLeuAspGluGluValAsnValIleProProHisThrProValArgThrValMetAsn 323
DE 1161 AAAGGATGATGAAAGAGGAAAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 1220
QY 424 ThrIleSerLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 343
DE 1221 ACTAATCAACAAATTAATGCTGATTAATTAATTAATTAATTAATTAATTAATTAAT 1280
QY 444 IleSerTyrPheAsnAsnGlyThrValAsnProLysGluSerIleLeuLysArgValLys 363
DE 1281 ATTTCGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1340
QY 464 AspIleSerTyrIleProLysGluLysAlaValGlyGlnGlyCysValGlu 383
DE 1441 GATGTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1400
QY 484 ThrIleSerLeuArgThrValArgLeuThrTyrAsnValMetGluSerMet 403
DE 1481 ATCCGAAATACAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440
QY 404 LeuLysSerGluGluArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAspAsn 423
DE 1461 CTTAAATCAGAAAGAAAGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1520
QY 424 IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValMetAlaThrTyrSerArg 443
DE 1521 AGTTTCATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1580
QY 444 SerThrSerGluAspLeuAspSerGlyThrAspLeuSerPheProTrpIleLeuAsnVal 463
DE 1581 AATAATACAGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1640
QY 464 IleAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPheIleLysAlaGluGly 483
DE 1641 CTTAAATTAAGGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1700
QY 484 AsnLeuThrArgGluMetIleCysIleuGluArgCysGluHisArgIleMetGluSer 503
DE 1701 AACTTGACAAAGAAAGAAATGATAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1760
QY 504 LeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGlnSerLysAspArgGlu 523
DE 1761 CTTGCAATGCTTCACATTCACCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1820
QY 524 GluProThrAspIleSerGluSerAlaCysProLeuAsnIleProLeuGlnAsnAsnHis 543
DE 1821 GAGCTTAAAGAACTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1877
QY 544 ThrAlaAlaAspMetIyrLeuSerProValArgSerProLysLysGlySerThrThr 563
DE 1878 ACTGAAAGAAATATATATATATATATATATATATATATATATATATATATATATAT 1937
QY 564 ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLys 583
DE 1938 CGTGTAATTTCTGCTGCAAAATACAGACACAAACAGCCCTCAGGCTTCCATACACAG 1997
QY 584 ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValIyrArgLeuAlaTyrLeu 603
DE 1998 CCATTGAAATCTACCTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2057
QY 604 ArgLeuAsnThrLeuCysGluArgLeuLeuSerGluHisProGluLeuGluHisIleIle 623
DE 2058 GAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2117
QY 624 TrpThrLeuPheGlnHisThrLeuGlnAsnGluTyrGluLeuMetArgAspArgLysLeu 643
DE 2118 TCCACATCTTCAGCAATACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2177
QY 644 ArgGluIleMetIleCysSerMetTyrIleGlyIleCysValLysAsnIleAspLeuLys 663
DE 2178 GAGCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2237
QY 664 PheLysIleIleValThrAlaTyrLysAspIleProHisAlaValGlnGluThrThrLys 683

```

```

DB 2238 TCCAAATCATCTCCATCCCTACAGCATCTCTCTCCAGCTGCCCCAGGACACCTTTAAA 2297
QY 684 ArgValLeuIleLysGluGluGlyTyrAspSerIleIleValPheTyrAsnSerValPhe 703
DB 2298 CCGTCTTTTCATCAGACAGACAGCAGTTTCATTCCTCATATAGTATTTCTATAAGTCGGTTTC 2357
QY 704 MetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProProThrLeuSer 723
DB 2358 ATGCAGAGACTAATAAACAATAATTTTATAGTATTAATTAATTAATTAATTAATTAAT 2417
QY 724 ProIleProHisIleProArgSerProTyrLysPheProSerSerProLeuArgIlePro 743
DB 2418 CCAATACCTACACATTCCTCAAGCTTTACAAAGTTCCTCAAGCTTTACAAAGCTTTACAA 2477
QY 744 GlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIleSerGluGlyLeuPro 763
DB 2478 GGAGTTTAATCATCTATATATATATATATATATATATATATATATATATATATATATAT 2537
QY 764 ThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyGluSerPheGly 783
DB 2538 ACACCAACAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2597
QY 784 ThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgValLeuLys 803
DB 2598 ACATCTGACAACTTCCAGAGAAATAAACCAGATGCTGCTAATAGTACAGACATCTCCAAA 2657
QY 804 ArgSerAlaGluGlySerAsnProProLysProLeuLysLysLeuAlaGlyAspIleGlu 823
DB 2658 AGAAGTTTGAAGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2717
QY 824 GlySerAspGluAlaAspGlySerLysHisLeuProGlyGluSerLysPheGlnLys 843
DB 2718 GGAGCTGATGAACAGAGATGGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 2777
QY 844 LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAspSerMet 863
DB 2778 CTGGCAGAAATCATTCCACTCCAGACAGAAATGCAAAAGCAGAGAGATTAATTAATTAAT 2837
QY 864 AspThrSerAsnLysGluGluLys 871
DB 2838 CATGCTCTCAAAACAGCAGCAAAAG 2861

```

Search completed: January 18, 2003, 22:32:21
Job time : 3916.94 secs

OR nucleic acid search, using SW model

Run on: January 16, 2003, 15:20:22 : Search time 203.96 seconds
(without alignments)
18469.506 Million cell updates/sec

File: us-09-026-459a-38
Perfect score: 6424
Sequence: 1 CGGGATGAGGAGGAAAAAC.....AAATGAGGATTATTCATAGT 3323

Scoring table: IDENTIV_NUC
Gapop 10.0 : Gapov 1.0

Searches: 16154056 steps, 809774376 residues
Total number of hits satisfying chosen parameters: 42408142

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

	EST:*
1:	em_esth1.*
2:	em_esth2.*
3:	em_esth3.*
4:	em_esth4.*
5:	em_esth5.*
6:	em_esth6.*
7:	em_esth7.*
8:	em_esth8.*
9:	em_esth9.*
10:	em_esth10.*
11:	em_esth11.*
12:	em_esth12.*
13:	em_esth13.*
14:	em_esth14.*
15:	em_esth15.*
16:	em_esth16.*
17:	em_esth17.*
18:	em_esth18.*
19:	em_esth19.*
20:	em_esth20.*
21:	em_esth21.*
22:	em_esth22.*
23:	em_esth23.*
24:	em_esth24.*
25:	em_esth25.*
26:	em_esth26.*
27:	em_esth27.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	910.2	27.4	1024	BM450031	BM450031 AGENCOURT
2	909.4	27.4	1051	BM451083	BM451083 AGENCOURT
3	856	25.8	1081	BM452221	BM452221 AGENCOURT
4	844.8	25.4	881	BM452275	BM452275 AGENCOURT
5	786	24.7	1001	BM475603	BM475603 AGENCOURT
6	752.8	22.7	1364	AK011246	AK011246 Mus muscu

7	752.4	22.6	1122	13	BM456881	BM456881 AGENCOURT
8	725.4	21.8	727	9	AL592831	AL592831 DKFZP3140
9	720.6	21.7	850	14	BM452227	BM452227 AGENCOURT
10	699.6	21.1	823	12	BM451061	BM451061 BM451061
11	698.2	21.0	792	9	AL563757	AL563757 AL563757
12	692.2	20.8	767	12	BM457056	BM457056 AGENCOURT
13	691.8	20.8	972	13	BM466194	BM466194 AGENCOURT
14	678	20.4	835	13	BM452724	BM452724 AGENCOURT
15	660	19.9	672	13	BM468359	BM468359 AGENCOURT
16	644.6	19.4	797	10	BM453928	BM453928 AGENCOURT
17	642.2	19.3	681	12	BM461619	BM461619 AGENCOURT
18	617	18.6	618	13	BM464009	BM464009 AGENCOURT
19	614	18.5	694	10	BM468246	BM468246 AGENCOURT
20	605.2	18.2	623	10	BM458181	BM458181 AGENCOURT
21	605	18.2	605	9	AL599633	AL599633 DKFZP3140
22	596.4	17.9	922	12	BM453533	BM453533 AGENCOURT
23	580	17.5	580	9	AL599395	AL599395 DKFZP3140
24	575.2	17.3	914	9	AL763411	AL763411 AGENCOURT
25	566.4	17.0	588	10	BM468244	BM468244 AGENCOURT
26	558	16.8	871	9	AL763405	AL763405 AGENCOURT
27	530.2	16.0	598	10	BM452767	BM452767 AGENCOURT
28	530	15.9	539	9	AL692790	AL692790 AGENCOURT
29	528	15.9	572	9	AL520941	AL520941 AGENCOURT
30	516.8	15.6	597	12	BM453246	BM453246 AGENCOURT
31	513.2	15.4	594	10	AL715533	AL715533 AGENCOURT
32	512.2	15.4	734	12	BM444857	BM444857 AGENCOURT
33	505.2	15.2	736	13	BM451568	BM451568 AGENCOURT
34	493.8	14.9	525	9	AL592555	AL592555 AGENCOURT
35	486.4	14.6	584	12	BM454050	BM454050 AGENCOURT
36	484.4	14.6	493	10	BM450287	BM450287 AGENCOURT
37	481.4	14.5	484	10	BM468095	BM468095 AGENCOURT
38	480.2	14.5	659	13	BM421116	BM421116 AGENCOURT
39	476.6	14.3	2083	13	BM456728	BM456728 AGENCOURT
40	474.4	14.3	734	13	BM452698	BM452698 AGENCOURT
41	459.2	13.8	457	12	BM434784	BM434784 AGENCOURT
42	457	13.8	457	9	AL093215	AL093215 AGENCOURT
43	457	13.8	732	9	AL598766	AL598766 DKFZP3140
44	454.2	13.7	869	12	BM455749	BM455749 AGENCOURT
45	447	13.5	447	9	AL082179	AL082179 AGENCOURT

ALIGNMENTS

RESULT 1	BM450031	1023 bp	mRNA	13004	EST 05 FEB 2002
LOCUS	AGENCOURT_6393544	NIH_MGC_72	Homo sapiens cDNA clone IMAGE:5528037		
DEFINITION	5' mRNA sequence.				
ACCESSION	BM450031				
VERSION	BM450031.1	GI:18494071			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1023)				
AUTHORS	NIH-MGC (http://mgi.nlm.nih.gov/)				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: quahs@mail.nih.gov Tissue procurement: ATCC/DCU/DBP cDNA library preparation: Life Technologies, Inc. DNA library arrayed by: the I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM12204 row 3 column 22 High quality sequence stop: 643. Location/Qualifiers 1. .1023				

FEATURES

Source


```

121 AATTTTAAATTCGAAAGCAACATTTTCAATGCTTTATTCGGCTGGCTCTT 180
124 GAGTTTAAATGTCATATAGCAGAAATATATCTCAAAATCTGATCTGCAAT 1308
141 GAGTTTAAATGTCATATAGCAGAAATATATCTCAAAATCTGATCTGCAAT 240
149 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 360
241 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 400
1469 ATCAAAAGTTTATCAAAAGCAGAAAGCACTTCACAAAGCAAAATGATAAAAGATTAGAA 1428
301 ATCAAAAGTTTATCAAAAGCAGAAAGCACTTCACAAAGCAAAATGATAAAAGATTAGAA 360
1429 GATGTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1488
461 GATGTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
1489 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
421 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
1549 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1608
481 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
1609 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 1668
641 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 600
1669 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1728
601 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
1729 AAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1788
661 AAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
1789 GATGTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
721 GATGTGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
1849 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 1908
781 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 840
1909 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1967
841 GTTATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
1968 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 2019
901 TGTGTTTGGCATGATTTGAAATGCTTAATTTTAAAGGCTTTGATTTTCAAAATG 953

```

```

RESULT 3
LOCUS B04.49221 1081 bp mRNA linear EST 24 MAY 2002
DEFINITION AGC000000.7766562 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015480
VERSION B04.49221
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC ht tp://mgi. nhl. gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: eudabs@mail.nih.gov

```

Tissue Procurement: APCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LAM3211 row: k column: 61
 High quality sequence stop: 732.

FEATURES

Location/Qualifiers
 1..1081
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6015480"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV_Script; Site: 1; Note:
 Site 2: Salt; Cloned unidirectionally; oligo dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 337 a 190 c 183 g 371 t

ORIGIN

Query Match 25.8%; Score 856; DB 14; Length 1081;
 Best Local Similarity 99.1%; Pred. No. 1-6e-157;
 Matches 892; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```

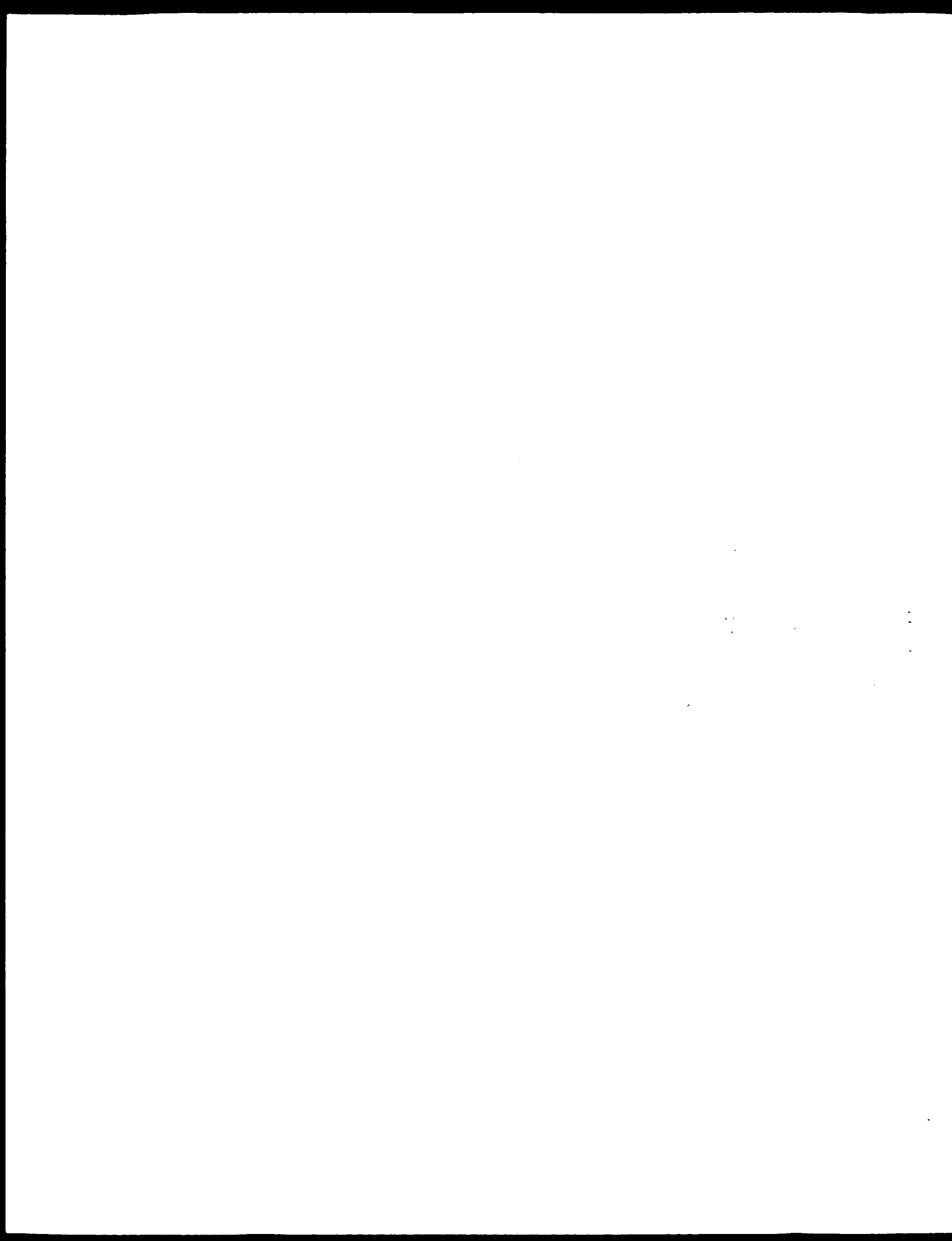
QY 2425 GAACACATGGAAGTAAACATCTCCAGCAGACATCCAAATTCACAGCAAAATGCAAA 2484
DB 1 GAACACATGGAAGTAAACATCTCCAGCAGACATCCAAATTCACAGCAAAATGCAAA 60
QY 2485 ATATTTCTATTTGAAATGCAAAATGCAAAATGCAAAATGCAAAATGCAAAATGCA 2544
DB 61 ATGACTTCTATCTCAACATGCAAAATGCAAAATGCAAAATGCAAAATGCAAAATGCA 120
QY 2545 AACACACAGAGAAATGAGATCTCAGATCTGCGGAGACATCTGATCTGATCTGATCT 2604
DB 121 AACACAGAGAGAAATGAGATCTCAGATCTGCGGAGACATCTGATCTGATCTGATCT 180
QY 2605 CATCTCTCTCAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2664
DB 181 CATCTCTCTCAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 240
QY 2665 ATATCTTCAATCTTTTGGGATATAAAATGTCAGATCTGATCTGATCTGATCTGATCT 2724
DB 241 ATATCTTCAATCTTTTGGGATATAAAATGTCAGATCTGATCTGATCTGATCTGATCT 300
QY 2725 AAGCCACATCAAAATGTCAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2784
DB 301 AAGCCACATCAAAATGTCAGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 460
QY 2785 TGGCATTTAAAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2844
DB 361 TGGCATTTAAAGTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 420
QY 2845 TGGATAGTAAATGCGGCTAGAGTGGGATCTGATCTGATCTGATCTGATCTGATCTGATCT 2904
DB 421 TGGATAGTAAATGCGGCTAGAGTGGGATCTGATCTGATCTGATCTGATCTGATCTGATCT 480
QY 2905 TGGCTTCTTTTGGACATAATAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2964
DB 481 TGGCTTCTTTTGGACATAATAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 2965 TTTTCTTAAATTAACATCAACACCTTACAAAATGTCATCTGATCTGATCTGATCTGATCTGATCT 3024
DB 541 TTTTCTTAAATTAACATCAACACCTTACAAAATGTCATCTGATCTGATCTGATCTGATCTGATCT 600
QY 3025 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3084
DB 601 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

```


[illegible]

2249	TTACTGTAAGATCAAGAAATCTTAATATCAATTTGGTGAATCATTTCTGGAGATCTTCGAAAT 2308	QY
60	TCACATCAAGATCAAGAAATCTTAATATCAATTTGGTGAATCATTTCTGGAGATCTTCGAAAT 120	DB
2309	TTCAAAAAATAAATCAATGCTATGTAACAGAGAACTGCTGCTCAAAAAGTGGTGAAG 2368	QY
121	TTCAAAAAATAAATCAATGCTATGTAACAGAGAACTGCTGCTCAAAAAGTGGTGAAG 180	DB
2369	GAAGCAACCTCTCTAAATCACTGAAAAAATACGCTTTGATATTGAAGCAACAGATGAAG 2428	QY
181	GAAGCAACCTCTCTAAATCACTGAAAAAATACGCTTTGATATTGAAGCAACAGATGAAG 240	DB
2429	CAGATGGAAGTAATATTTCTGAGAGAGTCCAAATTTCAAGCAAGCAACTGGTGAAGTAATGA 2488	QY
241	CAGATGGAAGTAATATTTCTGAGAGAGTCCAAATTTCAAGCAAGCAACTGGTGAAGTAATGA 300	DB
2489	CTTCTACTGAAACAGAAATGCAAAAAGTGAATGATAGCATGGATGATCTCAAGCA 2548	QY
301	CTTCTACTGAAACAGAAATGCAAAAAGTGAATGATAGCATGGATGATCTCAAGCA 360	DB
2549	AGAGACACAAAAGAGATCTTACGACCTTGGACACCTGGTACACCTCTGGATTCATT 2608	QY
361	AGAGACACAAAAGAGATCTTACGACCTTGGTGAACATGCTGGACACCTCTGGATTCATT 420	DB
2609	GTCTCTCAAGATGATGATATAACTTTTCCACGCTCTGTTATGGCCACATTTTAATAT 2668	QY
421	GTCTCTCAAGATGATGATATAACTTTTCCACGCTCTGTTATGGCCACATTTTAATAT 480	DB
2669	CTTCAAGCTCTTTTGTGATATAAAATGTGCATATGCAATTTGTTGGTGTATTCCTAAGC 2728	QY
481	CTTCAAGCTCTTTTGTGATATAAAATGTGCATATGCAATTTGTTGGTGTATTCCTAAGC 540	DB
2729	CAATTCAGAAAGTACGATATCTTATTTATACAGATTCAGAAATCTCTGCTAAATCTCTC 2787	QY
541	CAATTCAGAAAGTACGATATCTTATTTATACAGATTCAGAAATCTCTGCTAAATCTCTC 600	DB
2788	CATTTAAAAAATTTGTAGCAATATTTTCCCTCTTCCAAAGTAAAAATTTGCTGTGCTTTATGG 2847	QY
601	CATTTAAAAAATTTGTAGCAATATTTTCCCTCTTCCAAAGTAAAAATTTGCTGTGCTTTATGG 660	DB
2848	ATAGTAAAGTAATG 2859	QY
661	ATAGTAAAGTAATG 672	DB

Search completed: January 18, 2003, 04:08:48
Job time : 2939.95 secs



JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: ATCC

cDNA library preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: J1AM12263 Row: m Column: 18
High quality sequence stop: 737.

FEATURES
source

Location/Qualifiers
1..1051

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5550905"

/clone_lib="NIH_MGC_67"

/issue="viral-transformation"

/lab_host="DH10B (phage-resistant)"

/note="Organ cys. Vector: pMW-Sp-dT6, Site 2; Salt: Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Library constructed by Life

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

Technologies."

QY 544 ProLysLysLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAla 563
DB 544 CCAAGAGAAAAAGGTTCAACTACCGTGTAAATTTCTACTCAAAATGCAGACACACAGCA 603
QY 564 ThrSerAlaPheGlnThrGlnLysProLysSerThrSerLysSerLysThrLys 583
DB 604 AATCTCAAGTTTCAAGATCAAGATTAATTAATTAATTAATTAATTAATTAATTA 663
QY 584 LysValTyrArgGluAlaTyrLeuArgLysAsnThrLeuGluGluGluSerGlu 603
DB 664 AAGGTATATGGCTAGCCATATCTGGCTAAATACATTTGTGAACCCCTCTCTCTAC 723
QY 604 HisProGluLeuGluHisIleIleThrThrLeuPheGlnHisThrLeuGluAsnGluTyr 623
DB 724 CACCCAGAAATAGAACATATCAATCTGGACCCCTTTTCCAGCACACCCCTCCACAT 783
QY 624 GluLeuMetArgAspArgHisLeuAspGlnIleMetMetCysSerMetTyrGlyIleCys 643
DB 784 GAAATCATGAGAAAGAGCATTTTGAATATATATATATATATATATATATATATAT 843
QY 644 LysValLysAsnIleAspLeuLysPheLysIleIleValThrAlaLysLysAspLeu 863
DB 844 AAGGTGAAGATATAGACCTTAATATCAAAATCTCTGTACACACATACACAGATCTCC 903
QY 663 GluAlaValGlnGluThrPheLysArgValLeuLysLysGluGluTyrAspSer 682
DB 904 TCATGCTGTCAGGAGACATTCAAATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 963
QY 683 IleIleVal-----PheTyrAsnSerValPheMetGlnArgLeuLysThrAsnIle 699
DB 964 TATATAAGNATTCTTATAAATCTGGGATTTTATATATATATATATATATATATATAT 1020
RESULT 2
AK011246
LOCUS
DEFINITION
MUS musculus 10 days embryo whole body cDNA, RIKEN full-length
enriched library, clone:260017M22:retinoblastoma 1, full insert
sequence.
ACCESSION
AK011246
VERSION
AK011246.1 GI:12847243
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (strain-C57BL/6J) 10 days embryo cDNA to mRNA,
clone:260017M22.
ORGANISM
Mus musculus
Mukaryote; Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
Carninci, P. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE
20499374
PUBMED
11042159
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagakawa, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, P., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE
20530913

BASE COUNT 330 a 221 c 197 g 299 t 4 others

Alignment Scores:

Prod. No.: 2,64e-144 Length: 1051
Score: 1551.00 Matches: 312
Percent Similarity: 93.26% Conservat: 6
Best Local Similarity: 91.50% Mismatches: 18
Query Match: 34.83% Indels: 5
DB: 13 Gaps: 1

US-09-026-459a-43 (1-859) x BM561083 (1-1051)

QY 364 AlaValGlyGlnGlyCysValGluIleGlySerGlnArgTyrLysLeuGlyValArgLeu 383
DB 4 GCTGTGACACAGCTGT 63
QY 384 TTTTAT 403
DB 64 TATTACCAACTTCTGAAT 123
QY 404 PheSerLysLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGlu 423
DB 124 TTTATCAAACTTCTGAAT 183
QY 424 ValValMetAlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeu 443
DB 184 GHTCAAT 243
QY 444 SerPheProTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIle 463
DB 244 TCTTTCATGATGATTTTGAATATATATATATATATATATATATATATATATATATAT 483
QY 464 GluSerPheGlnLysAlaGluGluGluGluGluGluGluGluGluGluGluGluGlu 483
DB 304 GAAAGTTTATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAAT 363
QY 484 CysGluHisArgIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeu 503
DB 364 TGTGAACATCGAAT 423
QY 504 IleLysGlnSerLysAspArgGluGlyProThrAspHisLeuGluSerAlaCysProLeu 523
DB 424 ATTAACAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAAT 483
QY 524 AsnLeuProLeuGlnAsnAspHisThrAlaAlaAspMetTyrLeuSerProValArgSer 543
DB 484 AATCTGTCTGACAGAAATATCAATCTGATATATATATATATATATATATATATATATAT 543

with a Not I - oloigo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pU73 vector. Library is normalized. Library was constructed by Hento Soares and M. Fatima Ronaldo.

```

BASE COUNT      257 a 178 c 217 g 261 t 1 others
ORIGIN
Alignment Scores:      6,160-95      Length:      914
Pred. No.:      1056,50      Matches:      234
Percent Similarity:      82.14%      Conservative:      19
Best Local Similarity:      75.97%      Mismatches:      53
Query Match:      23,73%      Indels:      6
DB:      9      Gaps:      2

US 09 026 459a-43 (1-859) x AA764411 (1-914)

QY 267 PheLeuAspHisAspLysThrLeuGlnThrAspSerLeuAspSerPheGluThrGlnArg 286
DB 914 TTTTTCGACGACGACAAATAAGACGCTGCT--TCTTTAGGCGCTTTTAAAGCGCGGAG 857
QY 287 ThrProArgLysSerAsnLeuAspGluGluValAsnValIleProProHisIleProVal 306
DB 856 ACCGACGGAAACGACGCTTGT-GAAGCGGCAACAGAGGATTAAGGCTTAAAGGCTT 798
QY 407 ArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGln 326
DB 797 TGAGCGTCTTGGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG 738
QY 427 ProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIle 346
DB 747 GCGTCGACAAATTTGCTTTTTCGCTCTGCTAATATGGCGCGGGAAGGCAAGAAATTC 679
QY 347 LeuLysArgValLysAspIleGlyTyrIlePheGlySerGluGluValAlaGly 366
DB 678 CTAAAGCGGT-ANGAGGCTGGGACATGCTTAAAGAGAGATTTCTA---ACGCTGTGG 623
QY 367 GluGlyCysValGluIleGlySerGlnArgTyrIleLysGluGlyValArgLeuTyrArg 386
DB 622 CCAAGCGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563
QY 387 ValMetGluSerMetLeuLysSerGluGluGluArgLeuSerIleGluAsnPheSerLys 406
DB 562 GTCATGGAATCCATGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 503
QY 407 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuValValMet 426
DB 502 CTCTTAATGCAACATCTTCAATATGCTTTTACTGCTGCTGCTGCTGCTGCTGCTG 443
QY 427 AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro 446
DB 442 GCTATGCTATAGCAAGATATATGCGACGATTTGATTTGCGAAGACATTTGCTTCCG 383
QY 447 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPhe 466
DB 382 TGAATTCGCAACGCACTTAATTTAAAGCCCTTTCATTTTACAACTGATCAAGATTTT 323
QY 467 IleLysAlaGluThrLeuLeuThrArgGluMetIleLysHisLeuGlnArgCysGluHis 486
DB 422 ATCAAACTGCAAGCAACCTGCAAGCAAAATGATTAAGCAATTTTACAAAGATGTGAG 263
QY 487 ArgIleMetGluSerLeuAlaIlePheSerAspSerProLeuPheAspLeuLysGln 506
DB 262 CGAATCAAGAAATTCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 203
QY 507 SerLysAspArgGluGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuPro 526
DB 202 TCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 146
QY 527 LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 546
DB 146 CTGAGAGGTAAAGCATGAGGAGGATATGATATCTTCTCTCTCTCTCTCTCTCTCTCT 86

```

```

QY 547 LysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGluAlaThrSerAla 566
DB 85 AGAAGTCCACATACACAGCTGTAAATTCGCTGCGGAAATATACACACACACACAGCTCAGCC 26
QY 567 PheGluThrGluLysProLeuLys 574
DB 25 TTCATACACTGAGAAAGGCAATTGAAA 2

RESULT 10
AL563757/c
LOCUS      AL563757 LTI_NFL301_NB04 Homo sapiens cDNA (clone CS014047YA05) 4
DEFINITION prime, mRNA sequence.
ACCESSION      AL563757
VERSION        AL563757.1 GI:12913464
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 792)
AUTHORS      Li, W.H., Gruber, C., Jossee, J., and Pollayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
Genoscope - Centre National de Sequencage
HP 191 91006 kVRY codex - France
Email: seqref.genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS014047YA05"
/clone_lib="LTI_NFL301_NB04"
/sex="male"
/lissup_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      246 a 148 c 117 g 255 t 26 others
ORIGIN
Alignment Scores:      1,120-93      Length:      792
Pred. No.:      1043,00      Matches:      217
Percent Similarity:      84.53%      Conservative:      7
Best Local Similarity:      81.89%      Mismatches:      38
Query Match:      23,42%      Indels:      3
DB:      9      Gaps:      0

US-09-026-459a-43 (1-859) x AL563757 (1-792)

QY 171 LysThrAlaValIleProIleAsnGlySerProArgThrProArgArgGlyGluAsnArg 190
DB 791 AAAACAGCTTTTATACCACTAATGCTTACCTTCAGCAACATCCAGCGGACGTCAGACAG 732
QY 191 SerAlaArgIleAlaLysGlnLeuGluAsnAspThrArgIleIleGluValIleCysLys 210
DB 731 AGTGCAGCGGTAGCAGACACAAATAGAAATATGATACAGAAATATTGCAATTTCTCTAAA 672
QY 211 GluHisGluCysAsnIleAspGluValLysAsnValTyrIleLysAsnIleIleProLe 230
DB 671 GAAATGATGATGTATATATATATATATATATATATATATATATATATATATATATAT 612

```

QY 231 MetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGluAsnLeuSerLys 250
 DB 611 ATCAATTTCTTCCACTTCTTACATCTTATGACCTTCCAGAGCTTCAAAATTTCTTAA 552
 QY 251 ATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 270
 DB 551 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 492
 QY 271 AspLysThrLeuGluThrThrSerLysSerThrGluThrGluThrThrProGluLys 290
 DB 491 GATAAACCTCTTACAGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 432
 QY 291 SerAsnLeuAspGluGluValAsnValIleProProHisThrProValArgThrValMet 310
 DB 431 AGTAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 372
 QY 431 AsnThrIleGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 330
 DB 471 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
 QY 431 LeuLeuSerThrPheAsnAsnProLysThrValAsnProLysGluSerIleLeuLysArgVal 350
 DB 431 YGTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 252
 QY 451 LysAspIleGlyThrIlePheLysGluLysPheAlaLysAlaValGlyGluGlyCysVal 370
 DB 251 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 192
 QY 471 GluIleGlySerGluArgThrLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 390
 DB 191 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 132
 QY 491 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409
 DB 131 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 72
 QY 410 AspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuLeuValValMetAlaThrTyr 429
 DB 71 ACAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 12
 QY 430 SerArgSer 432
 DB 11 GCAGTAAT 3
 RESULT 11
 AL599633
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 FEATURES
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 605)
 KOEHLER,K., BEYER,A., MEWES,W., WEILL,B. and WIEMANN,S.
 EST (Kocher,K., Beyer,A., Mewes,H.W., Weill,B. and Wiemann,S.)
 Unpublished (1999)
 Contact: Koehler K
 MIPS
 Am Kleinferspitze 18a D-82152 Martinsried, Germany
 This is the 5' sequence of the clone insert
 Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charité,
 Berlin/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 No 5' sequence available.
 This clone (DKFZ313P092) is available at the EMBL in Berlin.
 Please contact the RZPD, Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY. Email: clone-arpd.de

FEATURES
 source
 Location/Qualifiers
 1..605
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="313 (synonym: hlcc2)"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Vector, pTriblox2; Site_1: SfilA; Site_2: SfilR;
 cDNA collection"
 BASE COUNT 228 a 32 c 105 g 179 t
 ORIGIN
 Alignment Scores:
 Prev. No.: 3 016-93 Length: 605
 Score: 1037.00 Matches: 201
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 23.29% Indels: 0
 DB: 9 Gaps: 0
 US-09-026-459a-43 (1-859) x AL599633 (1-605)
 QY 206 GluValLeuCysLysGluHisGluCysAsnIleAspGluValLysAsnValTyrPheLys 225
 DB 1 GAAGTTCTCTTAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 QY 226 AsnProIleProPheMetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluVal 245
 DB 61 AATTTTATATCTTTTATTAATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 120
 QY 246 GluAsnLeuSerLysArgTyrGluGluIleTyrLeuLysAsnLysAspLeuAspAlaArg 265
 DB 121 GAAATCTTTCTTAAAGATACGACGAAATTTATCTTAAATTAAGAATCTTAAATCTTAA 180
 QY 266 LeuPheLeuAspHisAspLysThrLeuGluThrAspSerIleAspSerPheGluThrGln 285
 DB 181 TTATTTTGGATCATGATAAACTCTTCTGACATATTTCTATGACATCTTTTCTAAATCT 240
 QY 286 ArgThrProArgLysSerAsnLeuAspGluGluValAsnValIleProProHisThrPro 305
 DB 241 AGAACACACAGGAAAGGTAAGCTTATGATGAGAGGTAAGTAATTCCTCTCTCTCTCTCT 300
 QY 306 ValArgThrValMetAsnThrIleGluGluGluGluGluGluGluGluGluGluGluGlu 325
 DB 301 GTTAGGACTGTTATGAACACTATCCCAACAAATTAATGATCAATTTTAAATTTCAAGTCA 360
 QY 326 GlnProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysLysSer 345
 DB 361 CAATTTTCAAAATCTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 420
 QY 346 IleLeuLysArgValLysAspIleGlyTyrIlePheGlySerGlyLysPheAlaLysVal 365
 DB 421 ATACTGAAAGAGTAAAGCATATAGATACATCTTTAAAGAGAAATTTCTTAAAGCTTIG 480
 QY 366 GlyGluGlyGlyValGluGluGlySerGluArgTyrLysLeuGlyValArgLysTyrTyr 385
 DB 481 GCAACAGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 386 ArgValMetGluSerMetLeuLysSerGluGluGluArgLeuSerIleGluAsnPheSer 405
 DB 541 GATGTAATGTAATCTATCTTAAATCAGAAAGAAATTTATTTATCTATCTATCTATCT 600
 QY 406 Lys 406
 DB 601 AAA 603
 RESULT 12
 HE082846/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 BE082846 1 61:8473151


```

/lab_host="DH10B"
/Note="Organ. Factors, Vector. pSP6K11, Site1: Not i,
Site 2: Sal I; Library constructed using SuperScript
Plasmid Library kit (Life Technologies). cDNA made by
oligo-dT priming. Size-selected by column fractionation;
average insert size 1.08 kb. Primary library,
unamplified."

```

```

BASE COUNT      234 a  93 c  109 g  186 t      1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3,950-91      Length:      623
Score:          1016.00      Matches:      201
Percent Similarity: 98.07%      Conservative:  2
Best Local Similarity: 97.10%      Mismatches:   3
Query Match:      22.82%      Indels:       1
DB:              10          Gaps:           0

```

```

US-09-026-459a-43 (1-859) x AW583181 (1-623)

```

```

QY 202 ThrAaTleTleTValLeuGysTysGtRhSsLysAsnLysValValLysAsn 221
|||||
Db 3 ACAGAAATATTCAAGTCTCTGAACACATCATTAATATATATAGATGAGTGAATAAT 62
|||||

QY 222 ValTyrPheLysAsnPhelIleProPheMetAsnSerLeuGlyLeuValThrScrAsnGl 241
|||||
Db 63 GTTATTTCAGAAATTTTATAGCTTTTATGGATTCTCTTGGACTTGTAAATCTAATGG 122
|||||

QY 241 yLeuProGluValGluAsnLeuSerLysArgTyrCluGluIleTyrLeuLysAsnLysAs 261
|||||
Db 123 ACHTTCAGAGGTTGAGAAATCTTTTAAACGATATGAAAGAAATTTATCTTAAATAAAGA 182
|||||

QY 261 pLeuAspAlaArgLeuPheLeuAspLysAspLysThrLeuGlnThrAspSerLleAspSe 281
|||||
Db 183 TCTAGATGCAAGATTATTTTGGATCATGATAAATCTTTAGACTGATTTATAGACAG 242
|||||

QY 281 rPheGlnThrGlnArgThrProArgLysSerAsnLeuAspGluCluValAsnValIlePr 301
|||||
Db 243 TTTTAAATACACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 302
|||||

QY 401 oProHisThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAs 321
|||||
Db 303 TCCACACATCCAGTACGACATGCTTATGACACATACCAACAAATTAAGATGATTTAAA 362
|||||

QY 421 nSerAlaSerAspGlnProSerGluAsnLeuLysSerTyrPheAsnAsnCysThrValAs 341
|||||
Db 363 TTTACAAATGATCAACCTTTACAAATCTGATTTCTATTTTAAACAAATTAAGATGAT 422
|||||

QY 341 nProGlnThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAs 361
|||||
Db 423 TTTACAAATGATCAACCTTTACAAATCTGATTTCTATTTTAAACAAATTAAGATGAT 482
|||||

QY 351 eAlaLysAlaValGlyGlnGlyCysValCluIleGlySerGlnArgTyrLysLeuGlyVa 381
|||||
Db 483 TGTAAAGCTCTGGACACAGGCTTCTCCAAATTTGATCAGCGGATACAACTTGGAGT 542
|||||

QY 481 lArgIleTyrTyrArgValMetLysSerMetLeuLysSerGluGluIleArgLeuSerTl 401
|||||
Db 543 TGTCTTGTATTAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAG 602
|||||

QY 401 eGlnAsnPhoSerLysLeu 407
|||||
Db 603 TCAAAATTTTACCAAACTT 621
|||||

```

```

RESULT 14
AA763485.1
LOCUS
DEFINITION
vW54h01 r1 Scores_mammay_21 and NMI MG Mus musculus cDNA clone
IMAGE:1247665.5 Similar to gp-M154h01 PETH08L8A10MA ASSOCIATED
PROTEIN (HUMAN); gp-M263q1 Mouse rEthinblastoma susceptibility
protein (MOUSE); mRNA sequence.
ACCESSION
AA763485
VERSION
AA763485.1 GI:2813232

```

KEYWORDS SOURCE ORGANISM

REFERENCE AUTHORS

TITLE JOURNAL COMMENT

```

EST.
house mouse.
Mus musculus.
Ekaryota: Metazoa: Chordata: Vertebrata, Euteleostomi:
Mammalia, Eutheria: Rodentia, Sciurognathi, Muridae: Murinae: Mus.
1 (bases 1 to 871)
Marr, M., Hillier, L., Allen, M., Howies, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kuvshinov, T., Lacey, M., Lee, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, P.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Maria M/Mouse tSI Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty free through LNCX. Contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI-661353
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 495.
Location/Qualifiers
1..871
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1247665"
/clone_lib="Soares_mammay_gland_NMI MG"
/sex="Female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT73 vector. Library is normalized. Library
was constructed by Benito Soares and M. Fatima Ronaldo."

```

FEATURES source

```

BASE COUNT      242 a  172 c  204 g  251 t      2 others
ORIGIN
Alignment Scores:
Pred. No.:      1,290-90      Length:      871
Score:          1013.00      Matches:      216
Percent Similarity: 90.34%      Conservative:  17
Best Local Similarity: 74.48%      Mismatches:   53
Query Match:      22.75%      Indels:       5
DB:              9          Gaps:           2
US-09-026-459a-43 (1-859) x AA763485 (1-871)
QY 287 ThrProArgLysSerAsnLeuAspGlnGlnValAsnValIleProHisThrProV 306
|||||
Db 860 AGGCGTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 804
|||||

QY 306 alArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspG 326
|||||
Db 803 TTGG-CCCGTTTTTCATCTACTCCACCATTAGGGGGCTTTTAAATTTCCCAAGTATCA 745
|||||

QY 326 InProSerGluAsnLeuLysSerTyrPheAsnAsnCysThrValAsnProLysGluSerI 346
|||||
Db 744 GAGCAACGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 685
|||||

QY 346 lLeuLysArgValLysAspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValG 366
|||||
Db 684 TTTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 625
|||||

QY 366 lYcGlnGlyCysValGlnIleGlySerGlnArgTyrLysGlnValArgLysThrVal 386
|||||

```


Search completed: January 19, 2003, 05:06:48
Job time : 1912.46 secs

Genome version 5.1.3
Copyright (c) 1993-2003 CompuGen Ltd.

em nucleotide - nucleotide search, using sw model

Run on: January 16, 2003, 15:20:22 : Search time 2771.81 Seconds
(without alignment)
18469.506 Million cell updates/sec

Title: us-09-026-459a-44

RefSeq score: 416

Sequence: 1 CAGGACAGGCGGCGCAAAAG.....AAATGACGATATATGATAGT 3161

Scoring table: IDENTITY NM

Gapop 10.0 / Gapext 1.0

Scoring: 16154066 seqs 400774376 residues

Total number of hits satisfying chosen parameters: 42404142

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listed first 4% summaries

Database: EST

- 1: em.estha.*
- 2: em.estha.*
- 3: em.estha.*
- 4: em.estha.*
- 5: em.estha.*
- 6: em.estha.*
- 7: em.estha.*
- 8: em.estha.*
- 9: em.estha.*
- 10: em.estha.*
- 11: em.estha.*
- 12: em.estha.*
- 13: em.estha.*
- 14: em.estha.*
- 15: em.estha.*
- 16: em.estha.*
- 17: em.estha.*
- 18: em.estha.*
- 19: em.estha.*
- 20: em.estha.*
- 21: em.estha.*
- 22: em.estha.*
- 23: em.estha.*
- 24: em.estha.*
- 25: em.estha.*
- 26: em.estha.*
- 27: em.estha.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query length	ID	Description
1	910.2	28.8	1523	13	BM450031 AGENCGURT
2	909.4	28.8	1051	14	BM561083 AGENCGURT
3	896	27.1	1081	14	BQ49221 AGENCGURT
4	752.4	23.8	1222	13	BM545881 AGENCGURT
5	743.2	23.5	881	14	BQ25275 AGENCGURT
6	725.4	22.9	727	9	AL597811 DKFZp313G

7	699.6	22.1	823	12	BC619661
8	692.2	21.9	767	12	BC570456
9	692.2	21.9	792	9	AL564757
10	691.8	21.9	972	13	BM466194
11	678	21.4	835	13	BM453724
12	660	20.9	672	13	BH68359
13	643.6	20.4	797	10	BP540278
14	642.2	20.3	691	12	BC261621
15	617	19.5	618	13	BM264009
16	614	19.4	694	10	BE082846
17	605.2	19.1	623	10	AW583181
18	605	19.1	605	9	AL599633
19	596.4	18.9	922	12	BC254543
20	580	18.3	580	9	AL599105
21	575.2	18.2	914	9	AA763411
22	566.4	17.9	588	10	AW368234
23	558	17.7	871	9	AA764885
24	530.2	16.8	598	10	AW582967
25	530	16.8	539	9	AL592790
26	528	16.7	572	9	AL120441
27	516.8	16.3	597	12	BE931236
28	513.2	16.2	594	10	AV715533
29	512.2	16.2	734	12	BE144857
30	505.2	16.0	736	14	BH151568
31	493.8	15.6	525	9	AA258255
32	486.6	15.4	1001	13	BM475603
33	486.4	15.4	584	12	BC149050
34	484.4	15.3	493	10	AW502887
35	481.4	15.2	484	10	BE168095
36	480.2	15.2	659	13	BH211116
37	476.6	15.1	2083	13	BM456728
38	474.4	15.0	734	13	BH525982
39	468	14.8	880	14	BQ222227
40	459.2	14.5	497	12	BE934784
41	457	14.5	457	9	A1093215
42	457	14.5	732	9	AL598766
43	454.2	14.4	869	12	BE65749
44	447	14.1	447	9	AL382179
45	446	14.1	453	10	AW504514

ALIGNMENTS

RESULT 1
BM450031
LOCUS
DEFINITION AGENCGURT.6393544 NIH_MGC_22 Homo sapiens cDNA clone IMAGE:5528047
5' mRNA sequence
ACCESSION BM450031
VERSION BM450031.1 GI:18494071
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Primates; Catarrhini; Hominoidea; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHORS NIH-MGC <http://mrc.nhl.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The L.M.A.G.E. Consortium (LML)
Clone Distribution: Agencourt Bioscience Corporation
found through the L.M.A.G.E. Consortium/LML at:
<http://image.llnl.gov>
Plate: LLML2504 row: d column: 22
High quality sequence stop: 643
Location/Qualifiers
SOURCE 1..1023

us-09-026-459a-44.rst

[illegible]

Db 121 AATTTAGTAAATTCAGAAACAAATTTTCAATGCTTTATGCGGTGGCTCTT 180
 QY 1087 GAGTTTAAATGATATATATAGAAATATATATAGAAATTTGATGGAACAT 1146
 Db 181 GAGTTTAAATGATATATATAGAAATATATATAGAAATTTGATGGAACAT 240
 QY 1147 TGGTTTAAATGATATATATAGAAATATATATAGAAATTTGATGGAACAT 1206
 Db 241 TGGTTTAAATGATATATATAGAAATATATATAGAAATTTGATGGAACAT 300
 QY 1207 ATGTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1256
 Db 401 ATGTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 360
 QY 1267 GATGTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1326
 Db 461 GATGTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 420
 QY 1427 CTATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1486
 Db 421 CTATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 480
 QY 1487 CTATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1446
 Db 481 CTATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 540
 QY 1447 TCTTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1506
 Db 541 TCTTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 600
 QY 1507 GATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1566
 Db 601 GATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 660
 QY 1567 AATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1626
 Db 661 AATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 720
 QY 1627 GATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1686
 Db 721 GATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 780
 QY 1687 TATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1746
 Db 781 TATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 840
 QY 1747 TATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1805
 Db 841 TATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 900
 QY 1806 TATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 1866
 Db 901 TATTAAGATTTTATTAAGAAATATATATAGAAATTTGATGGAACAT 960

RESULTS

4

6/4/9221

DEFINITION

6/4/9221

VERSION

6/4/9221

KEYWORDS

EST

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

6/4/9221 1081 bp mRNA linear EST 24 MAY 2002
 AGEN:0001_776562 NIH MGC_92 Homo sapiens cDNA clone IMAGE:6015480
 5' mRNA sequence
 6/4/9221
 6/4/9221.1 G:21179297
 EST
 human
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1081)
 NIH MGC <http://mgi.nhlbi.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Published (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: caples@mail.nih.gov

FEATURES
 source
 Location/Qualifiers
 1..1081

organism="Homo sapiens"
 /db_xref="Taxon:9606"
 /clone="IMAGE:6015480"
 /clone_lib="NIH MGC_92"
 /issue_type="Embryonal carcinoma, cell line"
 /lab_host="DH10B (phage resistant)"
 /note="organ: testis; Vector: pCMV SpeR6; Site 1: Not 1;
 Site 2: Salt: Cloned unidirectionally; oligo di primed;
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC library."

BASE COUNT 337 a 190 c 183 g 471 t

Query Match 27.1% Score 856; DB 14; Length 1081;
 Best Local Similarity 99.1% Pred. No. 4.4e 197;
 Matches 892; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

QY 2263 GAACACATGCAATTAATATGATGATATATATAGAAATTTGATGGAACAT 2322
 Db 1 GAACACATGCAATTAATATGATGATATATATAGAAATTTGATGGAACAT 60
 QY 2323 ATGACTTCTACTCGAACATGATGATGATATATAGAAATTTGATGGAACAT 2382
 Db 61 ATGACTTCTACTCGAACATGATGATGATATATAGAAATTTGATGGAACAT 120
 QY 2383 AACACATGCAATTAATATGATGATATATAGAAATTTGATGGAACAT 2442
 Db 121 AACACATGCAATTAATATGATGATATATAGAAATTTGATGGAACAT 180
 QY 2443 CATGCTCTCTCAGACATGATGATATATAGAAATTTGATGGAACAT 2502
 Db 181 CATGCTCTCTCAGACATGATGATATATAGAAATTTGATGGAACAT 240
 QY 2503 ATATCTTCAGCTCTTTTGGGATATATAGAAATTTGATGGAACAT 2562
 Db 241 ATATCTTCAGCTCTTTTGGGATATATAGAAATTTGATGGAACAT 300
 QY 2563 AACCCACTTGAATGTTAGTATGATATATAGAAATTTGATGGAACAT 2622
 Db 301 AACCCACTTGAATGTTAGTATGATATATAGAAATTTGATGGAACAT 360
 QY 2623 TGGCATTTAAAGATGATGATGATATATAGAAATTTGATGGAACAT 2682
 Db 361 TGGCATTTAAAGATGATGATGATATATAGAAATTTGATGGAACAT 420
 QY 2683 TGGTACTTAAAGATGATGATGATATATAGAAATTTGATGGAACAT 2742
 Db 421 TGGTACTTAAAGATGATGATGATATATAGAAATTTGATGGAACAT 480
 QY 2743 TGGCTCTCTCTCAGACATGATGATATATAGAAATTTGATGGAACAT 2802
 Db 481 TGGCTCTCTCTCAGACATGATGATATATAGAAATTTGATGGAACAT 540
 QY 2803 TTTTAAATTAAGATGATGATGATATATAGAAATTTGATGGAACAT 2862
 Db 541 TTTTAAATTAAGATGATGATGATATATAGAAATTTGATGGAACAT 600
 QY 2863 TTTGATGATGATGATGATGATATATAGAAATTTGATGGAACAT 2922
 Db 601 TTTGATGATGATGATGATGATATATAGAAATTTGATGGAACAT 660

Tissue procurement: AACC
 cDNA library preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 plate: L1AM1321 row: k column: 01
 High quality sequence stop: 732.

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02148
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@molbio.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center for information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@sim.wustl.edu)

Seq primer: 400P from Gibco
High quality sequence stop: 426.

FEATURES

Source

Location/Qualifiers

1..618

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="HBR85 islet"

/tissue_type="Purified pancreatic islet"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pBluescript SK(); Site_1:

Note: Site_2: XhoI; cDNA made by oligo-dT priming.

Size: selected on agarose gel. Average insert size 1kb. 5'

XhoI site was destroyed after directional cloning.

Amplified once. Contact information: Hiroshi Inoue, MD,

Metabolism Div. (Alan Permutt Lab); Washington University

School of Medicine, Box 8127, 660 South Euclid Ave., St.

Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:

314-362-1915, Fax: 314-747-2692."

BASE C UNIT 214 a 114 c 108 q 183 t

ORIGIN

Query Match.

Best local Similarity 100.0%; Score 617; DB 13; Length 618;

Mismatches 617; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2415 TGGCAAAATGATTTTATCTGAAATGCAATGCAAAATGAGAAATGATAGATGATGG 2374
DB 617 TGGCAAAATGATTTTATCTGAAATGCAATGCAAAATGAGAAATGATAGATGATGG 558
QY 2475 ATAGCTCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2434
DB 657 ATAGCTCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 498
QY 2445 CTGGGATTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2494
DB 497 CTGGGATTCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 438
QY 2495 CACATTTAATATCTGAGCTCTTTTGTGGATATATAAATGTGCAGATGCAATGTTTGGG 2554
DB 447 CACATTTAATATCTGAGCTCTTTTGTGGATATATAAATGTGCAGATGCAATGTTTGGG 378
QY 2555 TGATCTGAAGCACTTGAATGTTAGTCTATGTTTATATACAGATTGAATCTTGT 2614
DB 477 TGATCTGAAGCACTTGAATGTTAGTCTATGTTTATATACAGATTGAATCTTGT 318
QY 2615 GAAATCTGGCAATTAATAAGCTGATGAGATGTTCTCTCTCTCTCTCTCTCTCTCTCT 2674
DB 417 GAAATCTGGCAATTAATAAGCTGATGAGATGTTCTCTCTCTCTCTCTCTCTCTCTCT 258
QY 2675 GTGCTTTATATATATATATATATATATATATATATATATATATATATATATATAT 2734
DB 257 GTGCTTTATATATATATATATATATATATATATATATATATATATATATATATAT 198
QY 2745 GATATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2794
DB 147 GATATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 138
QY 2795 TAAGTAATTTTATATATATATATATATATATATATATATATATATATATATATATAT 2854
DB 147 TAAGTAATTTTATATATATATATATATATATATATATATATATATATATATATATAT 78
QY 2855 AATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2914
DB 77 AATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 18

```

QY 2915 GATATTTATTAGAAATTA 2931
DB 17 GATATTTATTAGAAATTA 1

Search completed: January 18, 2003, 04:10:13
Job time : 2798.89 secs

GenCore version 5.1.4
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein nucleic search, using frame_plus_p2n model

Run on: January 17, 2003, 18:27:24 : Search time 1763.22 seconds
(without alignments)
7320.587 Million cell updates/sec

Title: US-09-026-459A-45

Perfect score: 4144

Sequence: 1 MPKPTPRKTAATAAAAAAEP.....TPMKQPMNI-SMITSNREEK 797

Search not tables:

BLASTSUM62 Xgapop: 10.0 ; Xgapext: 0.5
Ygapop: 10.0 ; Ygapext: 0.5
Fgapop: 6.0 ; Fgapext: 7.0
Delop: 6.0 ; Delext: 7.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42308142

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Command line parameters:

Model: frame_plus_p2n.model -DEV-xlh
Q-z-p02: 1/208755, special/40500026459/runat_16012083_1521148_233222/aff-locury.fasta_2.133457
DB-EST: QMPT-fastap -SUFFIX-est -MINMATCH=0.1 -LOOPEL=0 -LOWPEXT=0
UNITS-bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
LOCALLEN=206 THRESHOLD=100 THRESHOLD=100 THRESHOLD=100 THRESHOLD=100 THRESHOLD=100 THRESHOLD=100
-GUTPMI-pro -WURD-EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=us09026459, us09026459, 16012083_1521148_233222 -MCP=6 -C=1000 3
NO XLDX: N2 MMAP -LARGEQUERY -NEG.SCORER=0 WAIT -LONGLOG -DEV-TIMEOUT=120
WARN TIMEOUT=40 THRESHOLD=10 XGAPOP=10 -XGAPEXT=0.5 -FAPOP=6 -FAPEXT=7
YGAPOP=10 YGAPEXT=0.5 -DELQ=6 -DELEXT=7

Database :

EST: *
1: cm estba: *
2: cm est hum: *
3: cm est in: *
4: cm estm: *
5: cm estm: *
6: cm estpl: *
7: cm estro: *
8: cm hre: *
9: qb est1: *
10: qb est2: *
11: qb hre: *
12: qb est3: *
13: qb est4: *
14: qb est5: *
15: cm est hum: *
16: cm estm: *
17: qb qss: *
18: cm qss hum: *
19: cm qss in: *
20: cm qss pl: *
21: cm qss vlt: *
22: cm qss hum: *
23: cm qss hum: *
24: cm qss mus: *
25: cm qss other: *
26: cm qss pro: *
27: cm qss rod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1551	37.4	1051	13	HM561083
2	1272	30.7	881	14	H2222275
3	1254	30.3	727	9	AL597811
4	1233.5	29.8	1464	11	AK911246
5	1208	29.2	835	13	BM453724
6	1111	26.8	1023	13	BM450031
7	1056.5	25.5	914	9	AA763411
8	1037	25.3	605	9	AA763411
9	1033	24.9	792	9	AL599643
10	1030	24.9	694	10	BE082846
11	1016	24.5	623	10	AW584181
12	1013	24.4	871	9	AA763485
13	1004	24.2	922	12	H2254543
14	978	23.6	588	10	AW468234
15	962	23.2	736	13	B1151768
16	959	23.1	594	10	AV715533
17	955	23.0	584	12	B6140050
18	947.5	22.9	734	12	BF144857
19	935	22.6	539	9	AL592790
20	908	21.9	797	10	H0549278
21	882	21.3	572	9	AL120941
22	879.5	21.2	734	13	B1526982
23	869.5	21.0	625	12	B6244388
24	858	20.7	597	12	H0931236
25	851	20.5	1001	13	HM475603
26	843	20.4	484	10	H0168095
27	834.5	20.1	646	13	BM490416
28	824	19.9	493	10	AW502887
29	815	19.7	622	9	AA072783
30	814	19.7	880	14	H2222227
31	783	18.9	2083	13	HM456728
32	751	18.1	453	10	AW503514
33	683	16.5	448	12	BF827047
34	681.5	16.4	799	14	CB2304
35	667	16.1	651	13	B1696841
36	659	15.9	380	10	AW516004
37	633	15.3	425	10	H225796
38	628	15.2	672	13	B1868359
39	626	15.1	521	12	BF466536
40	625.5	15.1	387	12	BF998490
41	613	14.8	415	9	AT745644
42	609	14.7	411	9	AT745643
43	594	14.3	1154	14	H0802746
44	593	14.3	580	9	AL599105
45	580	14.0	410	12	BE560103

ALIGNMENTS

RESULT 1
BM561083
LOCUS BM561083
DEFINITION AGNC00RT_6566133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550905
5' mRNA sequence.
VERSION HM561083
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS NIH-MGC <http://www.ncbi.nlm.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

BASE COUNT

ORIGIN

Alignment Scores:
 Pred. No.: 1,440-114 Length: 881
 Score: 1272.00 Matches: 254
 Percent Similarity: 98.08% Conservative: 1
 Best Local Similarity: 97.69% Mismatches: 4
 Query Match: 60.69% Indels: 2
 Gaps: 0

US-09-026-459A-45 (1-797) x H0220275 (1-881)

111 AlaValIleProLeuAsnGlySerProArgThrProArgArgGlyValAsnArgSerAla 130
 105 ACCTGTAATGCAATTAAGGTCACCTGCAATACCTGCAATACCTGCAATACCTGCAAT 164
 141 ArgLeuAlaGlyThrLeuValAsnGlySerProArgThrProArgArgGlyValAsn 150
 145 GAGATAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 224
 151 GlyValAsnIleAspGlyValIleValIleValIleValIleValIleValIleVal 170
 225 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 284
 171 SerLeuGlyLeuValThrSerAsnGlyLeuProGluValGluAsnLeuSerLysArgTyr 190
 285 TCTCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 344
 191 GluGluIleTyrLeuValAsnGlySerProArgThrProArgArgGlyValAsn 210
 345 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 404
 211 ThrLeuGluThrAspSerIleAspSerPheGluThrGluArgThrProArgLysSerAsn 230
 405 ACCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 464
 241 LeuAspGluIleValAsnValIleProProIleThrProValArgThrValMetAsnThr 250
 465 CTTCATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 524
 251 ThrGluGluLeuMetMetIleLeuValAsnSerAlaSerAspGluProSerGluAsnLeuIle 270
 525 ATCTGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 584
 271 SerTyrPheAsnAsnGlyThrValAsnProIleValIleValIleValIleValIle 290
 585 TCTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 644
 291 ThrGlyTyrIleThrLeuGlyGlyPheValGlyValGlyValGlyValGlyValGly 310
 645 ATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 704
 311 GlySerGluArgTyrLysGluGlyValArgGluTyrTyrArgValMetGluSerMetLeu 330
 705 GATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 764
 341 LysSerGluThrGluAlaGluSerIleGluAsnPheSerLysLeuLeuAsnAspAsnIle 350
 765 AAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 824
 351 PheIleMetSerLeuLeuAlaGlyAlaLeuGlyValValMetAlaThrTyrSer 368
 825 TCTATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 880

DEFINITION AL597811 727 bp mRNA linear EST 14-AUG-2001
 LOCUS DKE2P3160816.1 r1.613 (synonym: h1ec2) Homo sapiens cDNA clone
 ACCESSION AL597811
 VERSION AL597811.1 GI:15160502
 KEYWORDS EST.
 SOURCE human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 727)
 Poustka,A.; Wellenreuther,R.; Mewes,H.W.; Well,R. and Wiemann,S.).
 EST (Poustka,A.; Wellenreuther,R.; Mewes,H.W.; Well,R. and Wiemann,
 S.)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Poustka A.J.

Department: Lehrstuhl

Max-Planck-Institute for Molecular Genetics

Inmestrasse 73, 14195 Berlin, Germany

Tel.: +49-30-84131623

Fax: +49-30-84131128

Email: poustka@mpg-berlin.dahlem.mpg.de

This is the 5' sequence of the clone insert

Cloned from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email: S.Wiemann@dkfz-heidelberg.de;

Heidelberg/Germany) within the cDNA sequencing consortium of the

German Genome Project.

No sl sequence available.

This clone (DKE2P3160816) is available at the KZM in Berlin.

Please contact the RZPB: Ressourcenzentrum, Heidenberg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone-rzpb.de.

Location/Qualifiers

1..727

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="DKE2P3160816"

/clone_11b="413 (synonym: h1ec2)"

/dev_stage="adult"

/lab_host="DH10B"

/note="Vector: pTriplex2; Site_1: SmaI; Site_2: SmaI;

cDNA:collection"

BASE COUNT 238 a 161 c 130 g 198 t

ORIGIN

Alignment Scores:

Pred. No.: 6,190-114 Length: 727

Score: 1254.00 Matches: 241

Percent Similarity: 99.59% Conservative: 0

Best Local Similarity: 99.59% Mismatches: 1

Query Match: 40.26% Indels: 0

DB: 9 Gaps: 0

US-09-026-459A-45 (1-797) x AL597811 (1-727)

Q1 512 LysSerThrSerLeuSerLeuPheThrGlySerValIleArgLeuAlaTyrLeuArgLeu 531

DB 2 AAATCTACCT 61

Q1 532 AsnThrLeuGlySerGluArgLeuSerGluHisProGluLeuGluHisIleIleThrThr 551

DB 62 AATACACTTTTGAAAGCCCT 121

Q1 552 LeuPheGluHisThrLeuGluAsnGluThrGluLeuMetArgAspArgHisLeuAspGln 571

DB 122 CTCTTCAGACACACCT 181

Q1 572 IleMetMetGlySerMetTyrGlyIleGlyLysValIleAsnIleValIleValIle 591

DB 182 APTATGATGCT 241

Q1 592 IleIleValThrAlaTyrLysAspLeuProHisAlaValGluGluThrPheLysArgVal 611

DB 242 ATCATTTTAAACAGCATACAGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 401

Q1 612 LeuIleLysGluGluGluTyrAspSerIleIleValIleThrTyrAsnSerValPheMetGln 631

DB 302 TTGATCAAAAGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 461

Q1 632 ArgLeuLysThrAsnIleLeuGluThrValArgThrArgProThrLeuSerProIle 651


```

QY 505 PheGluThrGluLysProLeuLys 512
DB 25 LUCGATATTCAGAGCCATTCAGAA 2

RESULT 8
AL599633 605 bp mRNA linear EST 14-ADG-2001
DEFINITION DKF3313P092.1 313 (synonym: hlec2) Homo sapiens cDNA clone
AL599633 DKF3313P092.1 313 (synonym: hlec2) Homo sapiens cDNA clone
VERSION 1 (bases 1 to 605)
FEATURES
    EST (Koehler, K., Beyer, A., Mewes, W., Weill, B., and Wiemann, S.)
    TITLE EST (Koehler, K., Beyer, A., Mewes, W., Weill, B., and Wiemann, S.)
    JOURNAL Unpublished (1999)
    CONTACT Koehler K
    MISC Am Klepterspilz 18a D-82152 Martinsried, Germany
    This is the 5' sequence of the clone insert.
    clone from S. Wiemann, Molecular Genome Analysis, German Cancer
    Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
    sequenced by BMFZ (Biomedical Research Center at the Charité,
    Berlin/Germany) within the cDNA sequencing consortium of the German
    Genome Project.
    No 5' sequence available.
    This clone (DKF3313P092) is available at the RZPD in Berlin.
    Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
    Berlin-Charlottenburg, GERMANY; Email: clonacrzpd@rzd
    FEATURES
        1..605
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="DKF3313P092"
        /clone_lib="313 (synonym: hlec2)"
        /dev_stage="adult"
        /lab_host="DH10B"
        /note="Vectors: pTriplex2, site 1, SLIA, site 2, SLIIB,
        cDNA collection"
FAST 5' GNN 228 a 92 c 106 g 179 t
ORIGIN
Alignment Scores:
    Fred. No.: 1.63e-92 Length: 605
    Score: 1037.00 Matches: 201
    Percent Similarity: 100.00% Conservative: 0
    Best Local Similarity: 100.00% Mismatches: 0
    Query Match: 25.02% Indels: 0
    ID: 9 Gaps: 0
US 09-026-459A-45 (1-797) x AL599633 (1 605)
QY 144 GluValLeuGlyCysGluHisGluValLysAsnGluValLysAsnValThrLeuLys 163
DB 1 GAATTTCTGTGTAAACAAATCAATGTAATATAGATGAGTGCAAAATGTTTATTTCAAA 60
QY 164 AsnPheLeuProPheMetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluVal 183
DB 61 AATTTTATACCTTTTATGAAATCTCTTGGACCTTGAACATCAATGGACCTCCACAGGTT 120
QY 184 GluAsnLeuSerLysArgTyrGluGluLeuTyrLeuLysAsnLysAspLeuAspAlaArg 203
DB 121 GAAATCTTCTTAAAGATAGAGAAACAAATTTATCTTAAAAATAAAGATCTATATATGACAA 180
QY 204 LeuPheLeuAspHisAspLysThrLeuGluThrAspSerLysPheGluThrGln 223
DB 181 TTTATTTTGGATCATATATAAACTTTTAAATATATTATATATATATATATATATATATATAT 240

```

```

QY 224 ArgThrProArgLysSerAsnLeuAspGluGluValLysAsnValLysPheGluHisThrPro 243
DB 241 AGAACACACGAAAGAGTAACTTGATGAGAGGTCAGATTAATTCCTCCACACACTGTA 300
QY 244 ValArgThrValMetAsnThrTleGluGluLeuMetMetTleLeuAsnSerAlaSerAsp 263
DB 301 GTTAGGATGTTTATGAAACATCACTAAATTAATTAATTAATTAATTAATTAATTAATTA 360
QY 264 GluProSerGluAsnLeuLeuSerTyrPheAsnAspCysThrValAsnProGlySerGluSer 283
DB 361 CAATCTTATCAAAAATTCGATTTCTTATTTTAAACACTGCACAGTCAATCCAAAACAAAT 420
QY 284 TleLeuLysArgValLysAspLeuLeuTyrTlePheSerLysGluLysPheAlaLysAlaVal 303
DB 421 ATACTCAAAAGAGTGAAGGATATAGATACATCTTTAAAGACAAAATTTGCTAAAGCTGIG 480
QY 304 GlyGlnGlyCysValGluGluGlySerGluArgTyrTyrLysGluGlyValArgLeuTyrTyr 323
DB 481 GATAGAGGTTTGTCTGAATATTGGATCAAGAGCAATATAATTTGGAGATGGCTTGTTATTA 540
QY 324 ArgValMetGluSerMetLeuGlySerGluGluGluArgLeuSerLysGluAsnProGlySer 343
DB 541 CTAGTAACTCAATCTAATCTTAAATCAAGAGAGCAATATATCAATCAAAAATTTTAAAG 600
QY 344 Lys 344
DB 601 AAA 603

RESULT 9
AL563757/c
LOCUS AL563757 LTI_NFL001.NF-4 Homo sapiens cDNA clone CS00007YA05 3
DEFINITION AL563757 LTI_NFL001.NF-4 Homo sapiens cDNA clone CS00007YA05 3
ACCESSION AL563757
VERSION AL563757.1 GI:12913464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates, Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 792)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
        Genoscope - Centre National de Séquençage
        BP 191 91006 Evry cedex - France
        Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
FEATURES
    Location/Qualifiers
        1..792
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="CS00007YA05"
        /clone_lib="LTI_NFL001_NBC4"
        /sex="male"
        /tissue_type="neuroblastoma cells"
        /lab_host="DH10B"
        /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
        was primed with a NotI-oligo(dT) primer. Five prime end
        enriched, double-stranded cDNA was digested with Not I and
        cloned into the Not I and Eco RV sites of the pCMVSPORT 6
        vector. Library was normalized. Library was constructed
        by life technologies. Contact: Feng Zhang life
        Technologies, a division of Invitrogen 9800 Medical Center
        Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
        8371 Email: flianz@lifetech.com URL:
        http://fulllength.invitrogen.com"
BASE COUNT 245 a 148 c 117 g 255 t 26 others
ORIGIN
Alignment Scores:
    Fred. No.: 6.06e-92 Length: 792
    Score: 1033.00 Matches: 215

```


Genetic Screen Liver Vector pCMV SP6516, Site 1, No. 1
 Site 2: SalI: Cloned unidirectionally, oligo dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."

BASE COUNT
 ORIGIN

Alignment Scores:
 Pred. No.: 5,726-89 Length: 222
 Score: 1004.00 Matches: 224
 Percent Similarity: 90.51% Conservatives: 5
 Best Local Similarity: 88.54% Mismatches: 21
 Query Match: 24,23% Indels: 8
 DB: 12 Gaps: 0

US 09-026-459a-45 (1-797) x B0252543 (1-922)

QY 548 UctolrphthloupheGlnHrshrlouGlnAsnGln-TyrGlnuMotArGAsPAR 567
 DB 1 ATCAATGTGAGCCCTTTCAGAGCAACAGATGCA-CAATGAGATATGAATCATGAGAGAG 60
 QY 567 gHleGleAsnGlnLeuMeWetGserMeTyrGlyHleGlySerValLysAsnHleAs 587
 DB 61 GCATTC-GACCAAAATGATGATGTTCATGATGATGATGATGATGATGATGATGATGAT 119
 QY 597 pLeuLysPheLysHleValThrAlaTyrLysAspLeuProHleAlaValGlnGluTh 607
 DB 120 GCTTAAATTCAAATCAT-GTACACACATACACAGATCTTCCTCATGCTGTCACAGAC 178
 QY 607 rPheLysArValuHleLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 627
 DB 179 ATTCAAAAGTCT-TTGATTAAGAAAGAGATGATGATGATGATGATGATGATGATGATG 237
 QY 627 rValPheMetGlnArLeuLysThrAsnHleLeuGlnTyrAlaSerThrArqProProTh 647
 DB 238 GATCTGATGACACACACACAAACAAATAT-TTCAGATATGCTTCCACAGGCGCTTAC 296
 QY 647 rLeuSerProTleProHleLysPheProTyrLysPheProSerSerProLeuAr 667
 DB 247 ATTGCAAAATAGCTTCAATTCGTAAGAGGCTTTAAGATTTCTTATTAAGCTTACG 356
 QY 667 gHleGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 687
 DB 457 GATTCGAGAGAGAAATATATATTCACGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 416
 QY 687 yLeuProThProThLysMetThrProArgSerArgTleuValSerHleGlyGluSer 707
 DB 417 TCAGCAACACCAACAAATACATCCCAAGATCAAGAACTTAGTATCAATTCGCAATC 476
 QY 707 rPheGlyThrSerGlnLysPheGlnLysTleAsnGlnMeValLysAsnSerAspArV 727
 DB 477 ATTGCAAAATAGCTTCAATTCGTAAGAGGCTTTAAGATTTCTTATTAAGCTTACG 536
 QY 727 rPheLysArValuHleLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 747
 DB 537 GCTCAAAAGAGTGTGAG 595
 QY 747 rPheGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 766
 DB 596 TATTCAGCAACACCAACAAATACATCCCAAGATCAAGAACTTAGTATCAATTCGCAATC 655
 QY 767 GlnGlnLysLeuAlaGlnMeThrSerThrArqArqMetGlnGlnLysMetAsu 786
 DB 656 CAGAGCAAAATAG 716
 QY 787 AspSerMetAspThrSerAsnLysGluGluLys 797
 DB 716 GGAAGGAGAGATCTTCATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 748

RESULT 14

AW 468234

LOCUS

EST 04-FEB-2000

DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source

BASE COUNT
 ORIGIN

Alignment Scores:

Pred. No.: 1,126-86 Length: 588
 Score: 978.00 Matches: 191
 Percent Similarity: 97.45% Conservatives: 0
 Best Local Similarity: 97.45% Mismatches: 4
 Query Match: 23.60% Indels: 1
 DB: 10 Gaps: 0

US 09-026-459a-45 (1-797) x AW368234 (1-588)

QY 432 TrpLeuSerAspSerProLeuPheAspLeuLysGlnSerLysAspArgGlyGlyPro 451
 DB 3 TGNATGATGAGATCAATTTATTTCACTTATTAATTAATTAATTAATTAATTAATTAAT 42
 QY 452 ThrAspHleGlnGlnSerAlaGlySerProLeuAspLeuProLeuGlnAsnAsuHleThrAl 471
 DB 63 ACTGATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 422
 QY 471 gAlaAspMetTyrLeuSerProValArgSerProLysLysLysGlySerThrThrArqVa 491
 DB 123 AGCAGATATGATCTTCTTCGATGAGATCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAG 182
 QY 491 lAnSerThrAlaAsnAlaGlnThrGlnAlaThrSerAlaPheGlnThrGlnLysPheLe 511
 DB 183 AAATTCCTACTGCAAAATGCAGACACACACACACACACACACACACACACACACACAC 242
 QY 511 gLysSerThrSerLeuSerLeuPheTyrLysLysValTyrArqLeuAlaTyrLeuArqLe 531
 DB 243 GAAATATATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 502
 QY 531 uAsnThrLeuCysGluArqLeuLeuSerGlnHleProGlnLeuGluHleLleLleTrpH 551

CM368234-181000-024-c10 RT0181 Homo sapiens cDNA, mRNA sequence.
 AW368234
 AW368234.1 GI:5872884
 EST.
 human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

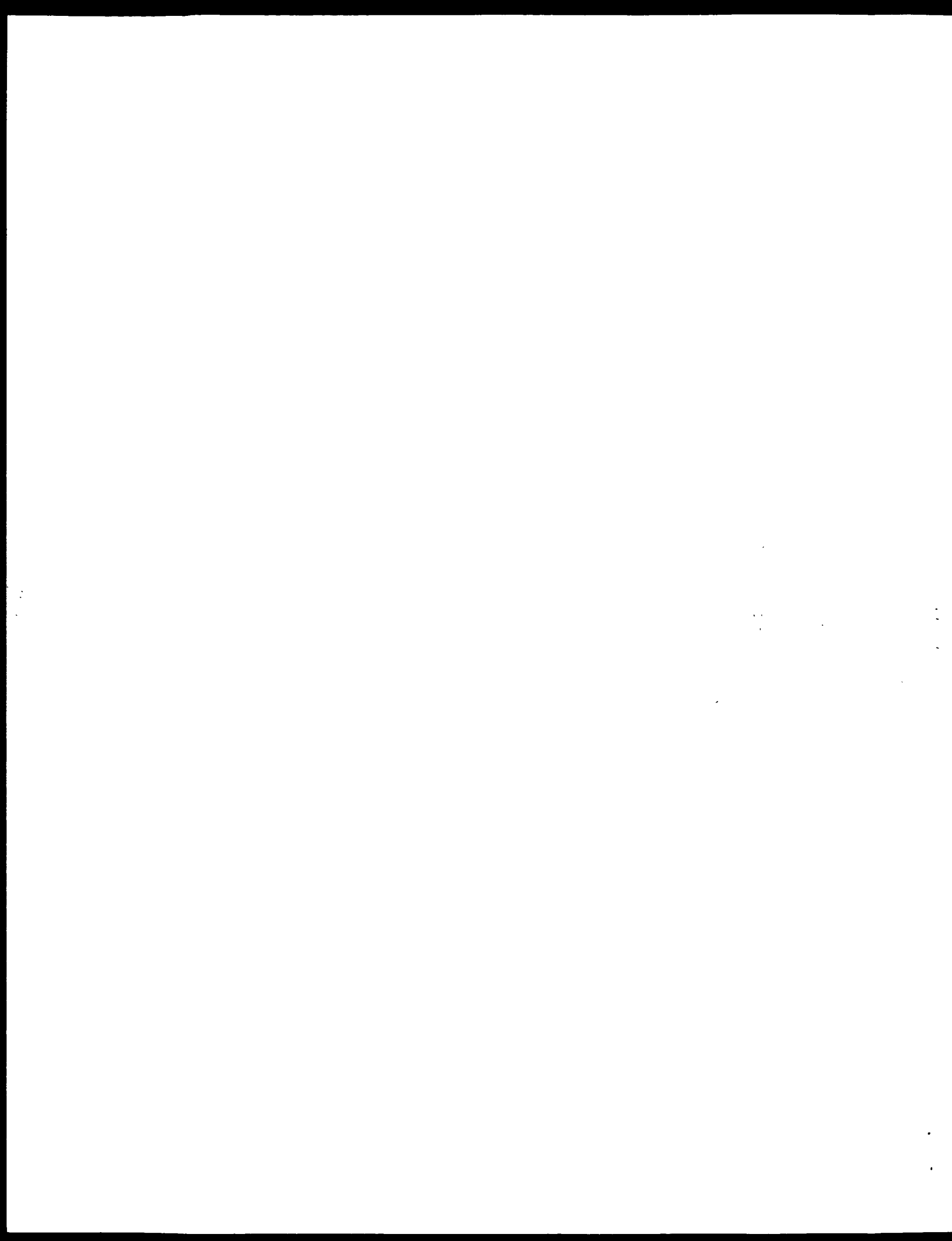
http://www.tudwig.org.br/PESTES.
 The FAFESP/LICR Human Cancer Genome Project
 Unpublished (1999)
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 105, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@tudwig.org.br
 This sequence was derived from the FAFESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.tudwig.org.br/scripts/gethtml2.pl?L1=CM368234-CM3-RT0181-
 181099-023-c10&t3=1999-10-18&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 559.

Location/Qualifiers

1..588
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /olecule="cDNA"
 /dev_stage="Adult"

Site 2: Small. A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pOC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions.

187 a 139 c 100 q 162 t



Q7	536	ProLysLysLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAla	555
Db	544	CCAAAGAAAAAGGTTAAATACGCTGTAAATCTACTGAAATGACACACACAAACA	603
Q7	556	maxSATApLapGlnThrGlnTyrPheLeuLysSerThrSerLeuSerLeuLys	575
Db	604	AACTATACCTTCTGAGACGAGAAAGGATGAAATATATCTCTTCTACTGTTTATAA	663
Q7	576	LysValTyrArgLeuAlaTyrLeuAlaArgAsnThrLeuGlySerAlaGluSerGln	595

Qy	596	HISPGGGLuLeuGLuGluHisLeuLeuLeuTETpThrLeuPheGLuHisThrLeuLeuGLuAspGLuTy	615
Db	724	CACCAGAAATTAGAACATATCATCTGGACCTTTTCCAGCACCCCTCCACAAATGACTAT	743
Qy	616	GGGAGGAGTATGATGATGGHISLeuAspGLuLeuMetMetCYSSTMetTyrGLyTyrLeuCYS	635
Db	764	GAATATATGAGATATATATATTTTATATATATATATATATATATATATATATATATATAT	783
Qy	636	LeuValLeuAspLeuLeuAspLeuLeuLeuValThrAlaTyrLeuPheAspLeuTyr	655
Db	844	AAAGTGAAGAGATATACACTTTTAAATTTTAAAGATGCTGCTGAGAGAGAGAGAGAGAG	903

```
Qy 655 -HisAlaValGlnGluThrPheTyrAsnValLeuIleLysLeuLeuLeuTyrAspSer 674
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 904 TCATCGTCAGCAGACATCAAACTGTGGTTGATGCAGAACAAGACGAGTAATGCC 963
Qz 675 HeHeLeval PheTyrAsnValPheMetGlnArgLeuGlyThrAsnIle 691
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 964 TATAAANATTGCTATAAATCGAGGCTTTATATGAAGTAAGTCAAAAACAAAA 1020
RESULT 2
AK011246
```

Accession	AK011246	Accession	AK011246
Definition	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2600017M22 ref:106181ama.1, full insert sequence.	Definition	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:2600017M22 ref:106181ama.1, full insert sequence.
Accession	AK011246	Accession	AK011246
Version	1	Version	1
Keywords	HTG; CAP trapper	Keywords	HTG; CAP trapper
Source	Mus musculus (Strain:57BL/6J) 10 days embryo cDNA to MEMA, clone:lib:PIKEN full-length enriched mouse cDNA library clone:2600017M22.	Source	Mus musculus (Strain:57BL/6J) 10 days embryo cDNA to MEMA, clone:lib:PIKEN full-length enriched mouse cDNA library clone:2600017M22.
Organism	Mus musculus	Organism	Mus musculus
Reference	1	Reference	1
Authors	Carninci,P. and Hayashizaki,Y.	Authors	Carninci,P. and Hayashizaki,Y.
Title	High efficiency full-length cDNA cloning	Title	High efficiency full-length cDNA cloning
Journal	Meth. Enzymol. 303, 19-44 (1999)	Journal	Meth. Enzymol. 303, 19-44 (1999)
Medline	99279253	Medline	99279253
PubMed	10349636	PubMed	10349636
Reference	2	Reference	2
Authors	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	Authors	Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Title	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	Title	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Journal	Genome Res. 10 (10), 1617-1630 (2000)	Journal	Genome Res. 10 (10), 1617-1630 (2000)
Medline	20499374	Medline	20499374
PubMed	11042159	PubMed	11042159
Reference	3	Reference	3
Authors	Shibata,K., Itoh,M., Aizawa,K., Nagata,S., Sasaki,N., Carninci,Konno,H., Akiyama,J., Nishi,K., Kitanai,T., Tashiro,H., Itoh,P., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,I., Harada,Y., Yamamoto,T., Matsumoto,H., Sakaguchi,S., Ikegami,T., Fushiwaki,Fujiwake,S., Inoue,K., Tozawa,Y., Izawa,M., Ohara,E., Watanabe,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Muramatsu,S., Kawaguchi,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.	Authors	Shibata,K., Itoh,M., Aizawa,K., Nagata,S., Sasaki,N., Carninci,Konno,H., Akiyama,J., Nishi,K., Kitanai,T., Tashiro,H., Itoh,P., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,I., Harada,Y., Yamamoto,T., Matsumoto,H., Sakaguchi,S., Ikegami,T., Fushiwaki,Fujiwake,S., Inoue,K., Tozawa,Y., Izawa,M., Ohara,E., Watanabe,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Muramatsu,S., Kawaguchi,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.
Title	RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multipillar sequencer	Title	RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multipillar sequencer
Journal	Genome Res. 10 (11), 1757-1771 (2000)	Journal	Genome Res. 10 (11), 1757-1771 (2000)
Medline	20530913	Medline	20530913


```

QY 673 AspSerIleIleValPheTyrAsnSerValPheMetGlnArqLeuLysThrAsnIleLeu 692
DB 674 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 667
QY 693 GlnTyrAlaSerThrArgProPro---ThrLeuSerProIleProHisIle-ProArgSe 711
DB 694 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 707
DB 698 CAGTATGTTTCACAGACGCGCGGAGGCTGTGTCACCAANTACATTCATTCCTCGAAG 727
QY 711 TProLysIle---PheProSerProLeu---ArgIlePro---GlyGlyAsn----- 726
DB 712 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 707
DB 728 GAGCGTTAAGATTTTCCTAGTTTCACGCTTTACGGAAATTCCTGCGAAGACCAATCC 787
QY 727 TleTyrIleSerProLeuLys 733
DB 728 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 809

RESULT 8
LOCUS BM450041 1023 bp mRNA linear EST 05-FEB-2002
DEFINITION ACEN-00RT-6493544 NIH-MGC-72 Homo sapiens cDNA clone IMAGE:5528037
5', mRNA sequence.
ACCESSION BM450041
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHOR NIG-MGC Project; NIH-Gov.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstraus@nhi.nih.gov
Tissue Procurement: AFCC/DCMP/DBP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plasmid: L14M12204 Row: 3 column: 22
High quality sequence stop: 643.
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="BM450041"
/issue_type="NIH-MGC-72"
/lab_host="DH10B (phage-resistant)"
/site "Orig. skin. Vector: pCMV-Sport6; Site 1: Not;
Site 2: Salt; Cloned unidirectionally. Primer: oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE 1 UNT 44 a 211 c 197 g 281 t
ORIGIN
Alignment Scores:
Pred No.: 118e-104 Length: 1023
Score: 111.00 Matches: 216
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.35% Indels: 0
DB: 13 Gaps: 0
US-09-026-459a-39 (1-851) x BM450041 (1-1023)
QY 696 LysValIleAsnIleAspLeuLysPheTyrIleValThrAlaTyrIleLysAspLeuPro 655
DB 697 TATATATATATATATATATATATATATATATATATATATATATATATATATAT 641
DB 2 AAAGTGAAGAAATATACACCTTAAATTCATATCTGTCTTACACCAATACACCAATCTCTCT 61
QY 656 HisAlaValGlnGluThrPheLysArqValLeuIleLysLysGluGluThrAspSerIle 675

```

```

DB 62 CATCGTTCAGAGACATCAAAAGTGTITGATCAAGAGAGAGATGATGATTT 121
QY 676 TleValPheTyrAsnSerValPheMetGlnArqLeuLysThrAsnIleLeuLysIleTyrAla 695
DB 122 ATATATCTTATATATATATATATATATATATATATATATATATATATATATAT 181
QY 696 SerThrArgProProThrLeuSerProIleProHisIleProArgSerProLysPhe 715
DB 182 TCCACGAGCGCCCTACTGTTCACCAATACCTTCACATTCCTCGAATCCCTTACAAGT 241
QY 716 ProSerProPro---ArgIlePheGlyGlyAsnIlePhe---GlyGlyAsnSerPro 745
DB 242 CCTAGTTCACGCTTACGGATTCCTGGAGGACACATCATATATTCACCTCGAAGAGTCCA 401
QY 736 TProLysIleSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 765
DB 302 TATATAATTTTCAGAGGCTGCTGCAACCAACCAACCAACCAACCAACCAACCAACCA 461
QY 756 ValSerIleGlyGlySerPheGlyThrSerGlyLysPheGlyLysIleAsnIleMetVal 775
DB 362 GTATCAATTCGCGAATCATTCGGGACTTTCAGAGGTTCCAGAAAATAAATCAGATGGTA 421
QY 776 CysAsnSerAspArqValLeuLysArgSerAlaGluGlySerAsnProProlLysProLeu 795
DB 422 TGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 481
QY 796 LysLysLeuArgPheAspIleGlyGlySerAspGlyLysAlaAspGlySerLysHisLeuPro 815
DB 482 AAAAAATAGCTTTGATATGAGAGATCAATGAGAGATCAATGAGAGATCAATGAGAGAT 541
QY 816 GlyGlySerLysPheGlyGlyLysLeuAlaGluMetThrSerThrArgThrArgMetGln 835
DB 542 GGACATTCCTCAATTTTCAGAGGATGAGAGATGAGAGATGAGAGATGAGAGATGAG 601
QY 836 LysGlnLysMetAsnAspSerMetAspThrSerAsnLysGluGluLys 851
DB 602 AAGCAGAAATATATATATATATATATATATATATATATATATATATATATATAT 649
RESULT 9
AA763411 914 bp mRNA linear EST 27-JAN-1998
LOCUS v54a04.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMAGE:1247598 5' similar to gb:M15400 RETINOBLASTOMA-ASSOCIATED
PROTEIN (HUMAN); gb:M26391 Mouse retinoblastoma susceptibility
protein (MOUSE);, mRNA sequence.
ACCESSION AA763411
VERSION AA763411.1 GI:2813158
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 914)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, I.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, P., Underwood, K., Moore, H.,
Theising, B., Wylie, T., Lennon, G., Soares, P., Wilson, R. and
Waterston, R.
The WashU-HUMI Mouse EST Project
Unpublished (1996)
Contact: Marra, M/Mouse EST project
WashU-HUMI Mouse EST Project
Washington University School of MedicineP
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through HUMI; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:661286
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 422.

```



```
801 pLleGluLysSerAspPheAlaAspGlySerIysHisLeuProGlyGluSerLysPhe 820
|||||
596 TATTAAGGATCAGATGAAAGGAGAAATCAAGTAAACATCCGCGCAGAGATCCCAATCT 655
|||||
821 GlnGlnLysLeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAsn 840
|||||
656 CAGGCAAGAAATGGGAAATGACTTATTGAAAAAGAAATGGCAAAAGGAAAGATCAAT 715
|||||
841 AspSerMetAspThrSerAsnLysGluLys 851
|||||
716 GGAAGCTGGTACCTCCACACAGGAGAAAGAAAG 748
|||||
```

Search completed: January 19, 2003, 05:06:23
Job time : 1895.76 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd

EM nucleic nucleic search, using SW model

Run on: January 16, 2003, 15:20:22 : Search time 404.47 Seconds
(without alignments)
18469.506 Million cell updates/sec

Files: US_09_026_459A_40

Perfect score: 3461
Sequence: 1 GGGGCAATGGGGGAAAAC.....AAATGACAAATATCACT 3461

Scoring table:

IDENTITY_NP*

Gapop 10.0 , gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42408142

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match ok

Maximum Match 100%

Listing first 45 summaries

Database :

```

1:  em_estb1a:
2:  em_estb1a:
3:  em_estb1a:
4:  em_estb1a:
5:  em_estb1a:
6:  em_estb1a:
7:  em_estb1a:
8:  em_estb1a:
9:  qb_est1a:
10: qb_est1a:
11: qb_est1a:
12: qb_est1a:
13: qb_est1a:
14: qb_est1a:
15: em_estb1a:
16: em_estb1a:
17: qb_est1a:
18: em_qss_hum:
19: em_qss_hum:
20: em_qss_hum:
21: em_qss_hum:
22: em_qss_hum:
23: em_qss_hum:
24: em_qss_hum:
25: em_qss_hum:
26: em_qss_hum:
27: em_qss_hum:

```

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.2	26.3	1023	13	BM450031 AGENCOURT
2	909.4	26.3	1051	13	BM561083
3	896	24.7	1081	14	BQ439221 AGENCOURT
4	844.8	24.4	881	14	BQ220275 AGENCOURT
5	790	22.9	1344	11	AK011246 Mus mus
6	786	22.7	1001	13	BM475603 AGENCOURT

7	752.4	21.7	1122	13	HM545881
8	725.4	21.0	727	9	AL597811
9	725.4	20.8	893	14	BM561083
10	699.6	20.2	824	12	BQ610661
11	698.2	20.2	792	9	AL563757
12	692.2	20.0	767	12	BM570456
13	691.8	20.0	972	13	BM466194
14	678	19.6	845	13	HM453724
15	660	19.1	672	13	B1868459
16	643.6	18.6	797	10	BM59278
17	642.3	18.6	681	12	BM616219
18	617	17.8	618	13	HM264009
19	614	17.7	694	10	HE082846
20	605.2	17.5	623	10	AW584181
21	605	17.5	605	9	AL599634
22	594.4	17.2	922	12	BM253543
23	580	16.8	580	9	AL599105
24	575.2	16.6	914	9	AA763411
25	566.4	16.4	588	10	AW368244
26	558	16.1	871	9	AA763485
27	530.2	15.3	598	10	AW582967
28	530	15.3	589	9	AL692790
29	528	15.3	572	9	AL139941
30	516.8	14.9	597	12	HE931236
31	513.2	14.8	594	10	AV715534
32	512.2	14.8	744	12	BF144857
33	505.2	14.6	746	13	B1151568
34	493.8	14.3	525	9	AA258256
35	486.4	14.1	584	12	BG149050
36	484.4	14.0	493	10	AW502887
37	481.4	13.9	484	10	HE168095
38	480.2	13.9	659	13	BF121116
39	476.6	13.8	2083	13	BM456728
40	474.4	13.7	744	13	B1526982
41	459.2	13.3	497	12	BE944784
42	457	13.2	457	9	AL094215
43	457	13.2	742	9	AL598766
44	454.2	13.1	869	12	HE865749
45	447	12.9	447	9	AL082179

ALIGNMENTS

RESULT 1
BM450031
LACTUS
DEFINITION AGENCOURT_6393544 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528037
5', mRNA sequence.

ACCESSION BM450031

VERSION BM450031.1 GI:18499071

KEYWORDS EST.

SOURCE Human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 1023)

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cvapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/UTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Cloning Strategy: Agencourt Bioscience Corporation

Cloning Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLNL2204 row 3 column 22

High quality sequence stop: 643

FEATURES
Location/Qualifiers
1..1023

[illegible]


```

QY 406 AAAATTGTTTAAATTAATAAATGCTGAAGTAAATGATATATATGATATATCA 3465
|||||
Db 718 TAAATTTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 776
|||||
QY 406 TATTAATTAATAAATGCTGAAGTAAATGATATATGATATATGATATATCA 3407
|||||
Db 777 TATTAATTAATAAATGCTGAAGTAAATGATATATGATATATGATATATCA 816
|||||

RESULT 11
AL563757 AL563757 L11_NFL001_NBC4 Homo sapiens cDNA clone EST 16-FEB-2001
LOCUS 792 bp mRNA linear EST 16-FEB-2001
DEFINITION prime, mRNA sequence.
VERSION AL563757.1 GI:12913464
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates, Catarrhini, Hominoidea; Homo.
1 (bases 1 to 792)
L11_NFL001_NBC4: Jesse, J. and Polys, B.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope, Centre National de Sequencage
BP 104 91006 Evry cedex - France
Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="L11_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="This NotI-Vesicle (NMV) cDNA library was primed with a NotI-Giga(CT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies. Contact : Feng Liang life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 246 a 148 c 117 g 255 t 26 other:s
BEGIN

Query Match: 20.2%, Score 698.2, DH 9, Length 792,
Best local similarity 93.6%, Pval No 2.1e-126,
Matches 722, Conservative 23, Mismatches 23, Indels 3, Gaps 2.

QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 489
|||||
Db 742 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 733
|||||
QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 749
|||||
Db 742 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 773
|||||
QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 809
|||||
Db 742 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 813
|||||
QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 869
|||||
Db 742 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 853
|||||
QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 929
|||||
Db 742 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 929
|||||

```

```

|||||
Db 552 ACATATAAATTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 493
|||||
QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 989
|||||
Db 492 TAAATTAATAAATGCTGAAGTAAATGATATATGATATATGATATATCA 433
|||||
QY 430 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1049
|||||
Db 432 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 873
|||||
QY 1050 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1109
|||||
Db 372 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 313
|||||
QY 1110 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1169
|||||
Db 312 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 253
|||||
QY 1170 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1229
|||||
Db 252 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 193
|||||
QY 1230 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1287
|||||
Db 192 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 133
|||||
QY 1288 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1347
|||||
Db 132 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 74
|||||
QY 1348 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 1397
|||||
Db 72 TAAACACCTGTTTAAATTAATAAATGCTGAAGTAAATGATATATGATATATCA 22
|||||

RESULT 12
BG570456 BG570456 767 bp mRNA linear EST 10-APR-2001
LOCUS 767 bp mRNA sequence.
DEFINITION BG570456
ACCESSION BG570456
VERSION BG570456.1 GI:13578109
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates, Catarrhini, Hominoidea; Homo.
1 (bases 1 to 767)
R111 M33 http://m33.bnl.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: rstra@bgl.nih.gov
Tissue: Eukaryotic, CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1568 row: h column: 03
High quality sequence stop: 726.
location/Qualifiers
1..767
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="R111_MGC_77"
/lab_host="DH10B (11 phage-resistant)"
/vector="Organ. lung. Vector. pDRM-LIB (Clontech); Site_1:
S111 (agccctcagcc); Site_2: S111 (agccctcagcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor

```


1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes the need for transparency and accountability in financial reporting.

2. The second part of the document outlines the various methods and techniques used to collect and analyze data. It includes a detailed description of the experimental procedures and the statistical analysis performed.

3. The third part of the document presents the results of the study. It includes a series of tables and graphs that illustrate the findings of the research. The data shows a clear trend of increasing activity over time, which is consistent with the hypothesis.

4. The fourth part of the document discusses the implications of the findings. It suggests that the results have significant implications for the field of research and may lead to further developments in the future.

5. The fifth part of the document concludes the study. It summarizes the main findings and provides a final statement on the importance of the research.

Department Lehrbach
Max-Planck-Institute for Molecular Genetics
Jahnstrasse 73, 14195 Berlin, Germany

Tel: +49-30-84131623
Fax: +49-30-84131128

Email: poustka@pimpq-berlin-dahlem.mpg.de

Clone from S. Wiemann, Molecular Genome Analysis

Research Center (DKFZ), Heidelberg, Germany, and the German Cancer Research Center (DKFZ) (German Cancer Research Center).

Heidelberg/Germany) within the cDNA sequence German Genome Project.

No si sequence available.
This clone (DKFZ031300816) is available at


```

67 739 AspSerIleValPheTyrAsnSerValPheMetGlnArdLeuLysThrAsnIleLeu 738
|||||
68 618 GATCTATATACATATATATATATATATATATATATATATATATATATATAT 667
69 749 GlyAlaSerThrArgProPhe ThrLeuSerProIleProHisIleProArgSe 757
|||||
70 768 CATATATATATATATATATATATATATATATATATATATATATATATATAT 727
71 767 ProIleTyrSerProIleProHisIleProArgSeThrAsnIleLeu 772
|||||
72 768 GATCTATATATATATATATATATATATATATATATATATATATATATATAT 787
73 773 IleTyrIleSerProIleTyrSerProIleProHisIleProArgSeThr 779
|||||
74 768 TATATATATATATATATATATATATATATATATATATATATATATATATAT 804

RESULT 8
BM450031 1023 bp mRNA linear EST 05-FEB-2002
ACFEPS-017-00044 NTH-MSC-01 nih.gov -DRA 01-02 IMA3F 053901
5' mRNA sequence.
BM450031
BM450031.1 GI:18499071
EST
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
1 (bases 1 to 1023)
NTH-MSC http://imgc.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rstraus@nhi.nih.gov
Tissue procurement: AICC/BCD/ZIP
cDNA Library Preparation: Life Technologies, Inc
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Amersham Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNI at:
http://image.llnl.gov
Plate: L14M1204 row: 4 column: 22
High quality sequence stop: 643.
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAF-5528037"
/clone_lib="NTH_MSC_72"
/tissue_type="melanotic melanoma"
/lab_host="JH10H (phage resistant)"
/vector="pGEM-3"
/site="NTH-MSC"
/site2="SalI, Cloned unidirectionally. Primer: oligo dT
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 44 a 211 c 197 g 281 t
ORGANISM

Alignment Scores:
Prod. No.: 1356 97 Length: 1023
Score: 1111.00 Matches: 216
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.99% Indels: 0
Gaps: 0

us-09-026-459a-41 (1-897) x BM450031 (1-1023)

67 739 AspSerIleValPheTyrAsnSerValPheMetGlnArdLeuLysThrAsnIleLeu 738
|||||
68 618 GATCTATATACATATATATATATATATATATATATATATATATATATATAT 667
69 749 GlyAlaSerThrArgProPhe ThrLeuSerProIleProHisIleProArgSe 757
|||||
70 768 CATATATATATATATATATATATATATATATATATATATATATATATATAT 727
71 767 ProIleTyrSerProIleProHisIleProArgSeThrAsnIleLeu 772
|||||
72 768 GATCTATATATATATATATATATATATATATATATATATATATATATATAT 787
73 773 IleTyrIleSerProIleTyrSerProIleProHisIleProArgSeThr 779
|||||
74 768 TATATATATATATATATATATATATATATATATATATATATATATATATAT 804

RESULT 9
AA763411 914 bp cDNA linear EST 27 JAN-1998
W54a04.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
IMA6-12476a 5' similar to 3' M15400 PBT18-RIAS2-MA-Associated
PROTEIN (HUMAN); gb:M26391 Mouse retinoblastoma susceptibility
protein (MCUR1);, mRNA sequence.
AA763411
AA763411.1 GI:2813158
EST
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 914)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubouche, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenger, K., Steptoe, M., Tan, F., Underwood, K., Moore, H.,
Theisinger, H., Wylie, T., Lennon, G., Soares, H., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8507, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty free through LLNI; contact the
IMAGE Consortium (image.llnl.gov) for further information.
MCI:651286
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 422.

```


603 1155 GAGGTTTTTGGATACATACAGACATTAAGGGGGGTTTTTAAATTCGCCAAGATACA 745
 QY 604 InProSerGluAsnLeuIleSerTyrPheAsnAspGlySerThrValAspProGlySerL 384
 DB 744 GCGATCGCAAAAATGCTGTTCCTGCTGGTTCATATTGTCGGGTACCTCCAAAGAACATC 685
 QY 684 LeuLeuArgArValLysAspIleGlyTyrIleGlySerGlyGlySerGlySerGlySer 404
 DB 684 GGTAAAGAGATTAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 625
 QY 404 TybLeuGlyValGlyIleGlySerGlySerGlySerGlySerGlySerGlySerGly 424
 DB 624 GCAAG 565
 QY 424 rValMetGluSerMetLeuLysSerGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 444
 DB 504 GTGGATTCGAAAGGATTAATTTAAAGGCTTTGATTTTAAAGATGATTTGATTTGAA 505
 QY 444 yLeuLeuAspAspAspAspAspAspAspAspAspAspAspAspAspAspAspAsp 464
 DB 504 CATTGCTAAATACAGAACATCTTATATGCTTACTGGGCTGCTGCTTCAAGATGTAA 445
 QY 464 rAlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhe 484
 DB 444 GAGTAAGATAGAGAGAGAGATACAGAGATCTGATCTGCAACACATCTGCTCTCC 385
 QY 484 rTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrTyrValIleGlySerP 504
 DB 684 GTGGATTCGAAAGGATTAATTTAAAGGCTTTGATTTTAAAGATGATTTGATTTGAA 325
 QY 504 LeuLeuLysAlaLeuLysAlaLeuLysAlaLeuLysAlaLeuLysAlaLeuLysAla 524
 DB 524 TTAATTAAG 495
 QY 524 IsAlaIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 544
 DB 204 ATCAAAATCAAGAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGATGATTTGAA 205
 QY 544 InSerLysAspArgProGlySerThrAspPheLeuGlySerAlaGlySerProLeuAsnLeuP 564
 DB 204 ATCAAAATCAAGAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGATGATTTGAA 148
 QY 564 rValLeuGlySerThrAlaLeuAspMetTyrLeuSerProValArgSerProLys 584
 DB 147 rTrpIleGlySerThrAlaLeuAspMetTyrLeuSerProValArgSerProLys 88
 QY 584 yLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGlyThrGlnAlaThrSerA 604
 DB 87 AAAGAACTTCACATACAG 28
 QY 604 LeuPheLeuThrGlnLysProLeuLys 612
 DB 27 rTrpIleGlySerThrAlaLeuAspMetTyrLeuSerProValArgSerProLys 2
 REFSeq 15
 LOCUS 602464278F1 NIH MGC 90 Homo sapiens cDNA clone IMAGE:4472611 5'
 DEFINITION 602464278F1 NIH MGC 90 Homo sapiens cDNA clone IMAGE:4472611 5'
 ACCESSION 602464278
 VERSION 602464278
 KEYWORDS EST
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 922)
 NIH MGC: <http://mimc.nhl.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: strausb@mit.edu
 Issue Procurement: A10C

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: the T.M.A.G.E. Consortium (LINC)
 cDNA Sequenced by: The Yale Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the T.M.A.G.E. Consortium/LINC at:
<http://lmage.linc.yoy>

Plate: L2AM10294 Row: 1 Column: 20
 High quality sequence stop: 633
 Location/Qualifiers
 1..922

FEATURES

Source
 1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4472611"
 /clone_1ib="NIH MGC 90"
 /issue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage resistant)"
 /note="Organ: Liver; Vector: pCMV SP-RP6; Site: 1; Not 1;
 Site_2: Salt; Cloned unidirectionally; cDNA di primed;
 Average insert size 1.7 Kb. Library enriched for
 full-length clones and constructed by life technologies
 Note: this is a NIH MGC library."

BASE COUNT 307 a 214 c 204 g 197 t

ORIGIN

Alignment Scores:
 Pred. No.: 3,130-87 Length: 922
 Score: 1034.90 Matches: 224
 Percent Similarity: 90.51% Conservatives: 5
 Best Local Similarity: 88.54% Mismatches: 21
 Query Match: 21.68% Indels: 8
 DNs: 12 Gaps: 0

US-09-026-459A-41 (1-897) x BG253543 (1-922)

QY 648 lleIleTrpThrLeuPheGlnHisThrLeuGlnAsnAla TyrGluLeuMetAlaAspAr 667
 DB 1 ATCACTCGACACGCTTTCAGACACACCTTCAGAACTCAGAACTCAGAACTCAGAACTCAG 60
 QY 647 gllsLeuAspGlnIleMetMetGlySerMetTyrGlyIleGlySerValIleAsnLeuAs 667
 DB 61 GCATTG-CAGCAAAATTATGATGCTTTCATGTATGGCATATG-AAAAGAGAAATATGA 119
 QY 687 pleuLysPheGlySerIleGlyValThrAlaGlyTyrGlyAspLeuGlnHisGlnGly 707
 DB 120 CTTTAAATTCAAAATCAT-GTAAACAGCAAT-AGAGATCTTCGATGCTTTCAGAACTCAG 178
 QY 707 rPheLysArgValLeuIleGlySerGlnGlyIleTyrAspSerIleGlyValPheTyrAsnLe 727
 DB 179 ATTCAAACTGCT-TCGATCAAAAGAGAGAGATGATGATGATGATGATGATGATGATGAT 247
 QY 727 rValPheMetGlnArgLeuLysThrAsnIleLeuGlnIleTyrAlaSerThrArgProGly 747
 DB 248 GGTCTTCATGACAGATGAGATGAGAAAT-AAAATAT-TCGATATGCTTCGATGATGATGAT 296
 QY 747 rLeuSerProIleProHisIleProArgSerProTyrTyrPheProSerSerProLeuAla 767
 DB 297 CTGTGCACCAATACATCATCTCTCGAAGCTCTTACAACTTACATGATGATGATGATGAT 456
 QY 767 gileProGlyGlyAsnIleTyrIleSerProLeuLysSerProTyrTyrLysIleSerIle 787
 DB 357 GATTTCGAGAGAGAGATCATATATTCACCTTCAGAGAGAGAGAGAGAGAGAGAGAGAG 416
 QY 787 yLeuProThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyGlyLeu 807
 DB 417 TCCTGCACCAATGAG 476
 QY 807 rPheGlyThrSerGlnLysPheGlnLysIleAsnGlyMetValGlySerGlyAspAlaIle 827
 DB 477 ATTCGGGACTTCTGAGAAATTCACAAATTAATCAATGATGATGATGATGATGATGATGAT 546
 QY 827 lleuLysArgSerAlaGlyGlySerAsnProProLysProLeuLysLysLeuArgPheAs 847
 DB 537 GCTCAAAAG 595

GenCore version 5.1.4
Copyright (c) 1993-2003 CompuGen Ltd.

CM protein nucleic search, using frame_plus.p2n model

Run on: January 17, 2003, 18:27:24 : Search time 2053.03 Seconds
(without alignments)
7420.587 Million cell updates/sec

Filter: US 09 026 459A 51

Perfect score: 4798
Sequence: 1 MPEKIKRKAIAAAAAAE.....IMCQKKNISMTSNKEEK 928

Scoring table:
BLOSUM62 Xgapop 10.0 : Xgapext 0.5
Ygapop 10.0 : Ygapext 0.5
Fgapop 6.0 : Fgapext 7.0
Dgapop 6.0 : Dgapext 7.0

Searches: 16154066 seqs, 8047743376 residues

Total number of hits satisfying chosen parameters: 32408142

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Command line parameters:

MODEL=frame.p2n.model -DEV-klh
Q-Z=us09026459a51 -SUFFIX=1 -MINMATCH=0.1 -LOOPCL=0 -LXOPCL=0
UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
DOALIGN=200 -THR_SCORE=100 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MDDP=LOCAL
-GAPEXT=10 -GAPOP=10 -HEAP=125-500 -MINLEN=0 -MAXLEN=200000000
USER US09026459a51 1 18175 -format_152148_23322 -NCPU=6 -ICPU=4
NO XLPXY NO MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120
WARN TIMEOUT 40 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
YGAPOP=10 YGAPEXT=0.5 DELOP=6 -DELEXT=7

Database :

EST: *
1: cm.est.bm: *
2: cm.est.bm: *
3: cm.est.in: *
4: cm.est.in: *
5: cm.est.in: *
6: cm.est.in: *
7: cm.est.in: *
8: cm.est.in: *
9: qb.est.1: *
10: qb.est.2: *
11: qb.est.3: *
12: qb.est.4: *
13: qb.est.5: *
14: qb.est.6: *
15: cm.est.in: *
16: cm.est.in: *
17: qb.qss: *
18: cm.qss.bm: *
19: cm.qss.bm: *
20: cm.qss.in: *
21: cm.qss.in: *
22: cm.qss.in: *
23: cm.qss.in: *
24: cm.qss.in: *
25: cm.qss.in: *
26: cm.qss.in: *
27: cm.qss.in: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1883	97.2	1364	11	AK011246 Mus musculus
2	1551	32.3	1051	13	BM561084 AGENCOBRT
3	1448	40.2	881	14	HQ229275 AGENCOBRT
4	1339	27.9	1001	13	BM475603 AGENCOBRT
5	1254	26.1	727	9	AL597811 UKF2p116
6	1246.5	25.8	880	14	BQ229277 AGENCOBRT
7	1208	25.2	835	13	BM454724 AGENCOBRT
8	1111	23.2	1024	13	BM450041 AGENCOBRT
9	1091	22.7	832	9	AI646338
c 10	1056.5	22.0	914	9	AA764411
c 11	1043	21.7	792	9	AI564757
c 12	1037	21.6	605	9	AL599643
c 13	1030	21.5	694	10	BE082846
c 14	1016	21.2	624	10	AW584381
c 15	1013	21.1	871	9	AA764485
16	1004	20.9	922	12	BQ229275
17	978	20.4	588	10	AW368244
18	963.5	20.1	709	9	AI646097
19	962	20.0	736	13	B1151568
20	959	20.0	594	10	AV715533
21	955	19.9	584	12	HG149050
22	947.5	19.7	734	12	BF144857
23	935	19.5	534	9	AI592790
24	908	18.9	797	10	DE539278
25	903	18.8	615	9	AA959836
26	890	18.5	1154	14	BM62746
27	882	18.4	572	9	AL120941
28	879.5	18.3	744	13	H1526382
29	869.5	18.1	525	12	HG244088
c 30	858	17.9	557	12	BE941246
c 31	844	17.6	446	10	BE168095
c 32	843.5	17.4	646	13	BM490416
33	824	17.2	494	10	AA502887
34	815.5	17.0	522	9	AA072780
35	813	16.9	545	10	HG692240
36	792	16.5	643	13	B1553937
37	784	16.3	2083	13	BM456728
38	751	15.7	453	10	AA503514
39	704	14.7	529	9	AA959709
40	693.5	14.5	472	9	AA955457
c 41	690	14.4	462	9	AI606529
c 42	684	14.2	448	12	BF927047
c 43	681.5	14.2	799	14	CB2404
44	681	14.2	514	12	BF418684
45	667	13.9	651	13	B1696841

ALIGNMENTS

RESULT 1
AK011246
LOCUS Mus musculus 1364 bp mRNA library RIKEN full length
DEFINITION Mus musculus 10 days embryo whole body cDNA, RIKEN full length
enriched library, clone:260017M22:refinedblastom1, full insert
sequence:
ACCESSION AK011246
VERSION AK011246.1 GI:12847243
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) 10 days embryo cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library
clone:260017M22.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

100

ier: 55-11-2704922
 FAX: 55-11-2707001
 Email: asimerson@uic.edu
 This sequence was derived from the FAPESP/LICP Human Genome
 Project. This entry can be seen in the following URL
 (<http://www.ncbi.nlm.nih.gov/nucl/2704922>)
 295 314 33643-2000 02 2181-1
 Seq primer: puc 18 forward
 High quality sequence start: 58
 High quality sequence stop: 669.

FEATURES

BASE COUN		211 a	129 c	127 g	1 others
<p>stringency conditions.</p>					
<p>0051H</p>					
<p>Alignment Scores:</p>					
Pred. No.:		1,42e '92	Length:		694
Score:		1040.00	Matches:		207
Percent Similarity:		96.33%	Conservative:		3
Best Local Similarity:		94.95%	Mismatches:		7
Query Match:		21.47%	Indels:		2
Gap:		10	Gaps:		0
<p>HS-09 226.456A-51 (1-928) x BE082846 (1-694)</p>					


```
QY 555 isAqalleMetGluSerLeuAlaTrpLeuSerAspSerLeuLeuPheAspGluHisLysG 575
1b 264 ATGAAATCAATGGAAATCCCTTCCAGAGCTTTTCAGATTCACCTTTATTTGATCTCATTAAGC 205
QY 575 InSerLysAspArgGluGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuP 595
1b 284 ASTCCAAAGATGCAAGAGACCT---GATAACCTTGAAGCTTTCTTTCTCTCTTCAACCTGC 148
27 595 rGleuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValAlaLysSerProLysL 615
1b 147 CTCTCAGAGGTAAAGCAATAGTGGAGAGATATGTAATCTTCTCTCTAAGATCTGCAAGA 88
QY 615 ySLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerA 635
1b 87 AAATAACTTCCACTACACCTGTAATAATTCCTGCTCAAAATACAGAGACATACAGGCTCAG 28
QY 635 laPheGlnThrGlnLysProLeuLys 643
1b 27 CCTTCATATAGTCAAGAGCATTTGAAA 2
```

Search completed: January 19, 2003, 05:07:40
Job time : 2687.11 secs

GenBank version 5.1.13
Copyright (c) 1993 - 2003 Computer Ltd.

EM nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:20:22, Search time 316.42 seconds
(without alignments)
18469,506 Million cell updates/sec

Hit: 08-09-026-459a-50

Score: 4554
Sequence: 1 GCGTCATGCGGCGCAAC.....AAATGAGGATTATGATAGT 4554

Scoring table: IDENTITY NC
Gap: 10.0, GapExt: 1.0

Searched: 16154666 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42308132

Minimum DB seq length: 0

Maximum DB seq length: 20000000

Post processing: Minimum Match: 0%

Maximum Match: 100%

List of first 45 summaries

Database: EST:

1: em_estb1

2: em_estb1

3: em_estb1

4: em_estb1

5: em_estb1

6: em_estb1

7: em_estb1

8: em_estb1

9: em_estb1

10: em_estb1

11: em_estb1

12: em_estb1

13: em_estb1

14: em_estb1

15: em_estb1

16: em_estb1

17: em_estb1

18: em_estb1

19: em_estb1

20: em_estb1

21: em_estb1

22: em_estb1

23: em_estb1

24: em_estb1

25: em_estb1

26: em_estb1

27: em_estb1

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Length	ID	Description
1	968.4	27.2	164	11 AK011246	Mus muscu
2	910.2	25.6	1023	14 BM450031	BM450031 AGENCGURT
3	909.4	25.6	1051	14 BM450031	BM450031 AGENCGURT
4	856	24.1	1081	14 BM450031	BM450031 AGENCGURT
5	844.8	23.9	981	14 BM450031	BM450031 AGENCGURT
6	786	22.1	1001	14 BM450031	BM450031 AGENCGURT

RESULT 1	AK011246	1464 bp	mus mus	19 JAN 2002
LOCUS	Mus musculus 10 days embryo whole body cDNA, RIKEN full length enriched library, clone:2609017M22,refinedblastoma 1, full insert sequence.			
DEFINITION	AK011246.1 GI:12847243			
ACCESSION	AK011246.1			
VERSION	AK011246.1			
KEYWORDS	HTC; CAP trapper.			
SOURCE	Mus musculus (strain:G78L6J) 10 days embryo cDNA to mRNA, clone:lib-RIKEN full-length enriched mouse cDNA library, clone:2609017M22.			
ORGANISM	Mus musculus			
REFERENCE	1 Fukuyama, M. et al. Chordata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Mus. 1			
AUTHORS	Carninci, P. and Hayashizaki, Y.			
TITLE	High efficiency full-length cDNA cloning			
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)			
MEDLINE	99279253			
PubMed	10349636			
REFERENCE	2			
AUTHORS	Carninci, P., Shibata, Y., Hayashizaki, Y., Saitoh, K., Itoh, M., Kono, H., Ozaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes			
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)			
MEDLINE	20499374			
PubMed	11042159			

AUTHORS: Poustka, A., Wellenreuther, R., Mewes, H.W., Weill, B., and Wiemann, S.)
 TITLE: EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Weill, B., and Wiemann, S.)

GENERAL COMMENT:
 Published (1999)
 Department: Poustka A.J.
 Max-Planck-Institut für Molekulare Genetik
 Ihnestr. 73, 14195 Berlin, Germany
 Tel.: +49-30-84131623
 Fax: +49-30-84131128
 Email: poustka@mpg-berlin-dahlem.mpg.de
 This is the 5' sequence of the clone insert
 clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ). Email: s.wiemann@dkfz-heidelberg.de;
 sequenced by DKFZ (German Cancer Research Center,
 Heidelberg, Germany) within the cDNA sequencing consortium of the
 German Genome Project.
 No sl sequence available.
 This clone (DKFZ3130816) is available at the RZPD in Berlin.
 Please compare this paper, Poustka et al., Nucleic Acids Res. 27, 14059
 Berlin-Charlottenberg, GERMANY; Email: clonereg@dkfz.de

FEATURES
 source
 1..727
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="DKFZ3130816"
 /clone_lib="313 (synonym: hicc2)"
 /dev_stage="adult"
 /lab_host="DH10b"
 /seq_org="cDNA" /prim1="P2, 5'-TTC-1 SFL1A, 5'-TTC-2 SFL1B;
 cDNA collection"
 BASE COUNT 238 a 161 c 130 g 198 t

Query Match 20.4%; Score 725.4; DP 9; Length 727;
 Best Local Similarity 99.9%; Pred. No. 8.4e-192;
 Matches 726; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1332 GAATGATGCTCTTTACTGTTTATATAAAAGTATGCTGATGCTATGCTGCT 1901
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 1 GAATGATGCTCTTTACTGTTTATATAAAAGTATGCTGATGCTATGCTGCT 60
 QY 1992 AAAATACCTTCTGCAAGTATGCTGATGCTGATGCTGATGCTGATGCT 2051
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 61 AAAATACCTTCTGCAAGTATGCTGATGCTGATGCTGATGCTGATGCTGAT 120
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 1352 AATTTTGAAGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 2111
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 121 GCTTTCTGACACACCTGCAAGTATGCTGATGCTGATGCTGATGCTGATGCT 180
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2112 AATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2171
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 181 AATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 240
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2172 AATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2231
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 241 AATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 400
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2232 TTTGATCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 2291
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 301 TTTGATCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 360
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2292 GAAAT 2351
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 361 GAAAT 420
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2352 AATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2411
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 421 AATTTGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 480
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2412 GAAAT 2471
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 481 GAAAT 540
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT

QY 2472 AACAAAAAGATCCAGATCAAGAAAGCTTAACTAACAATGGAGAAACATGGAGATTC 2531
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 541 AACAAAAAGATCAAGATCAAGAAAGCTTAACTAACAATGGAGAAACATGGAGATTC 600
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2532 TGAGAGTTCCAGAAAAATATATATATATATATATATATATATATATATATATAT 2591
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 601 TGAGAGTTCCAGAAAAATATATATATATATATATATATATATATATATATATAT 660
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2592 TGCTAAGAAAGAAATATATATATATATATATATATATATATATATATATAT 2651
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 661 TGCTAAGAAAGAAATATATATATATATATATATATATATATATATATATAT 720
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 2652 ACAATGAA 2658
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 721 ACAATGAA 727

RESULT 9
 HQ222227
 Locus
 DEFINITION AGENT:HQ222227.1 C:20403627
 5' mRNA sequence.
 ACCESSION HQ222227
 VERSION HQ222227.1 C:20403627
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 880)
 AUTHORS NIH-MGC http://mgi.nhlbi.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rstraus@nih.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLRC)
 DNA Distribution by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLRC at:
 http://image.llrc.org
 Plate: LHAM3279 row: i column: 20
 High quality sequence stop: 606.

BASE COUNT 321 a 148 c 152 g 259 t
 ORIGIN
 Location/Qualifiers
 1..880
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6041563"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal
 /lab_host="PH10B (phage-resistant)"
 /note="Organ: Testis; Vector: pCMV-SV0R15; Site: LLRC; North
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH-MGC Library."

Query Match 20.3%; Score 720.6; DP 14; Length 880;
 Best Local Similarity 98.7%; Pred. No. 7.2e-131;
 Matches 737; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 475 GATTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 534
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 1 GATTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 60
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 QY 535 GAAACAGTTCGATATGATATGATATGATATGATATGATATGATATGATATGAT 594
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT
 DB 61 GAAACAGTTCGATATGATATGATATGATATGATATGATATGATATGATATGAT 120
 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT
 TT

sequences: 5' CAGCGGCAATTAAGGTC 4' and 4' adaptor sequence:
5' ATTCTACAGGCGAGCGGCGATG-dt(30)NN 3' (where N = A,
C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC library."

```

BASE COUNT      233 a   125 c   140 g   279 t
CGTGTG

Query Match      19.5%  Score 692.8; DB 12; Length 767;
Best Local Similarity 97.1%  Pred. No. 2,66 125;
Matches 74; Conservative 0; Mismatches 18; Indels 4; Gaps 4;

QY 2762 GATGATATAGTCTAAAGAAAGAAATAGAGATCTGAGAGCTTGCTGCAATCTGTG 2821
DB 1 GATGATATAGTCTAAAGAAAGAAATAGAGATCTGAGAGCTTGCTGCAATCTGTG 60

QY 2822 TATACCTTCTGCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2881
DB 1 TATACCTTCTGCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120

QY 2882 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2941
DB 1 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

QY 2942 TTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3001
DB 1 TTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

QY 3002 TTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3061
DB 1 TTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400

QY 3062 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3121
DB 1 ATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360

QY 3122 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3181
DB 1 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 3182 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3241
DB 1 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 3242 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3300
DB 1 TCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY 3300 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3360
DB 1 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 3360 TTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
DB 1 TTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 659

QY 3420 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3480
DB 1 AATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718

QY 3480 TTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3540
DB 1 TTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767

QY 3540 TTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600
DB 1 TTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 767

```

PE3001114
 ESTS RM466194 972 bp mRNA linear EST 05-FEB-2002
 DEFINITION ACEN000001_045673 NIH MGC_92 Homo sapiens cDNA clone IMAGE:5576939
 ACCESSION RM466194
 5' mRNA sequence.

RM466194.1 GI:18515236

EST.
 human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 (bases 1 to 972)

Authors

Title

Journal

Comment

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: rgs@nsl.nih.gov

Tissue Procurement: AICC

cDNA Library Preparation: Life Technologies, Inc.

Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Library Sequenced by: Agencourt Bioscience Corporation

Library Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL2429 row: j column: 12

High quality sequence stop: 621.

Location/Qualifiers

1..972

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="NIH MGC_92"

/issue_type="embryonal carcinoma, cell line"

/lab_host="pH10R (phage-resistant)"

/note="origin: testis; vector: pCMV Sport6; Site 1: Not1;

Site 2: SalI; cloned unidirectionally; oligo dt primed.

Average insert size 2.5 kb. Library enriched for

full length clones and constructed by Life Technologies.

Note: this is a NIH MGC Library."

2 others

BASE COUNT 294 a 184 c 169 g 423 t

ORIGIN

Query Match 19.5% Score 691.8; DB 13; Length 972;

Best Local Similarity 96.7% Pred. No. 3,16 125;

Matches 728; Conservative 0; Mismatches 19; Indels 6; Gaps 2;

QY 2698 CAGCAGAAATGGCAGAAATGACTTCTACGCAACACAGCAAGCAAAATCAAT 2757

DB 1 CAGCAGAAATGGCAGAAATGACTTCTACGCAACACAGCAAGCAAAATCAAT 60

QY 2758 GATAGCATGATATCTCAAAACAGGAAAGAAATAGATCTGAGATCTGAGTAC 2817

DB 61 GATAGCATGATATCTCAAAACAGGAAAGAAATAGATCTGAGATCTGAGTAC 120

QY 2818 TCTCTACACCTCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2877

DB 121 TCTCTACACCTCTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180

QY 2878 TGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2937

DB 181 TGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

QY 2938 ATTGTTGGGTGATCTCTAGGCACTTGAATTTTATTTATTTATTTATTTATTT 2997

DB 241 ATTGTTGGGTGATCTCTAGGCACTTGAATTTTATTTATTTATTTATTTATTT 400

QY 2998 AAAATCTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3057

DB 401 AAAATCTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 460

QY 3058 TAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3117

DB 361 TAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 3118 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3177

DB 421 AGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480


```

/Ab xref "taxon:9606"
/clone="IMAGE:5402622"
/clone_id="NIH_MG7_90"
/issue_type="adenocarcinoma, cell line"
/lab_host="PHU8 (phage-resistant)"
/notes="organ: liver; Vector: pCMV-SPORT6; Site_1: Nott;
Site_2: Salt; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full length clones and constructed by life technologies.
Note: this is a NIH MCC library."
BASE COUNT      228 a   124 c   144 g   186 t
CGTCTN

```

```

Query Match      18.6%   Score 660; DB 13; Length 672;
Best local Similarity 99.9%   Prod. No. 5,3e 119;
Matches 671; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

97 2420 ATATTTCAGCTGGAGATGCAATATAATAATTTGAGAGGTGTGCAACAGCAACAAAA 2479
100 1 ATATTTCAGCTGGAGATGCAATATAATAATTTGAGAGGTGTGCAACAGCAACAAAA 60

97 2480 GCACTGCAAGCAAGCAAACTTACAGCAATAGGCAATCATTCGGGACCTGCGAGAGT 2539
100 1 GCACTGCAAGCAAGCAAACTTACAGCAATAGGCAATCATTCGGGACCTGCGAGAGT 120

97 2540 TCCAGCAATATAATACATGATGATGTAACAGGAGGTGTGCTCAAGCAAGGTGCTGAAG 2599
100 1 TCCAGCAATATAATACATGATGATGTAACAGGAGGTGTGCTCAAGCAAGGTGCTGAAG 180

97 2600 GAGGAGGCTGCTTAAGCACTGCAAAAAATCTAGTTTGATATTGAAGGATCAGATGAAG 2659
100 1 GAGGAGGCTGCTTAAGCACTGCAAAAAATCTAGTTTGATATTGAAGGATCAGATGAAG 240

97 2660 CATAGCAAGCAAAATCTGCTAGGCAAGCTGCAAAATTCAGCAGCAAACTGGCAGCAATCA 2719
100 1 CATAGCAAGCAAAATCTGCTAGGCAAGCTGCAAAATTCAGCAGCAAACTGGCAGCAATCA 300

97 2720 CTTCCTATCTGCAACCAATGCAAAAATGAGAAATGAATGATAGCATGCTCAAAACA 2779
100 1 CTTCCTATCTGCAACCAATGCAAAAATGAGAAATGAATGATAGCATGCTCAAAACA 360

97 2780 AGCAAGCAAAATGAGCAATCTGCAAGCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCT 2839
100 1 AGCAAGCAAAATGAGCAATCTGCAAGCTGCGGCAATGCTGCTGCTGCTGCTGCTGCTGCT 420

97 2840 GCTCTCTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2899
100 1 GCTCTCTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

97 2900 CTTCAGCTCTTTTGGGATATAAAAATGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2959
100 1 CTTCAGCTCTTTTGGGATATAAAAATGCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

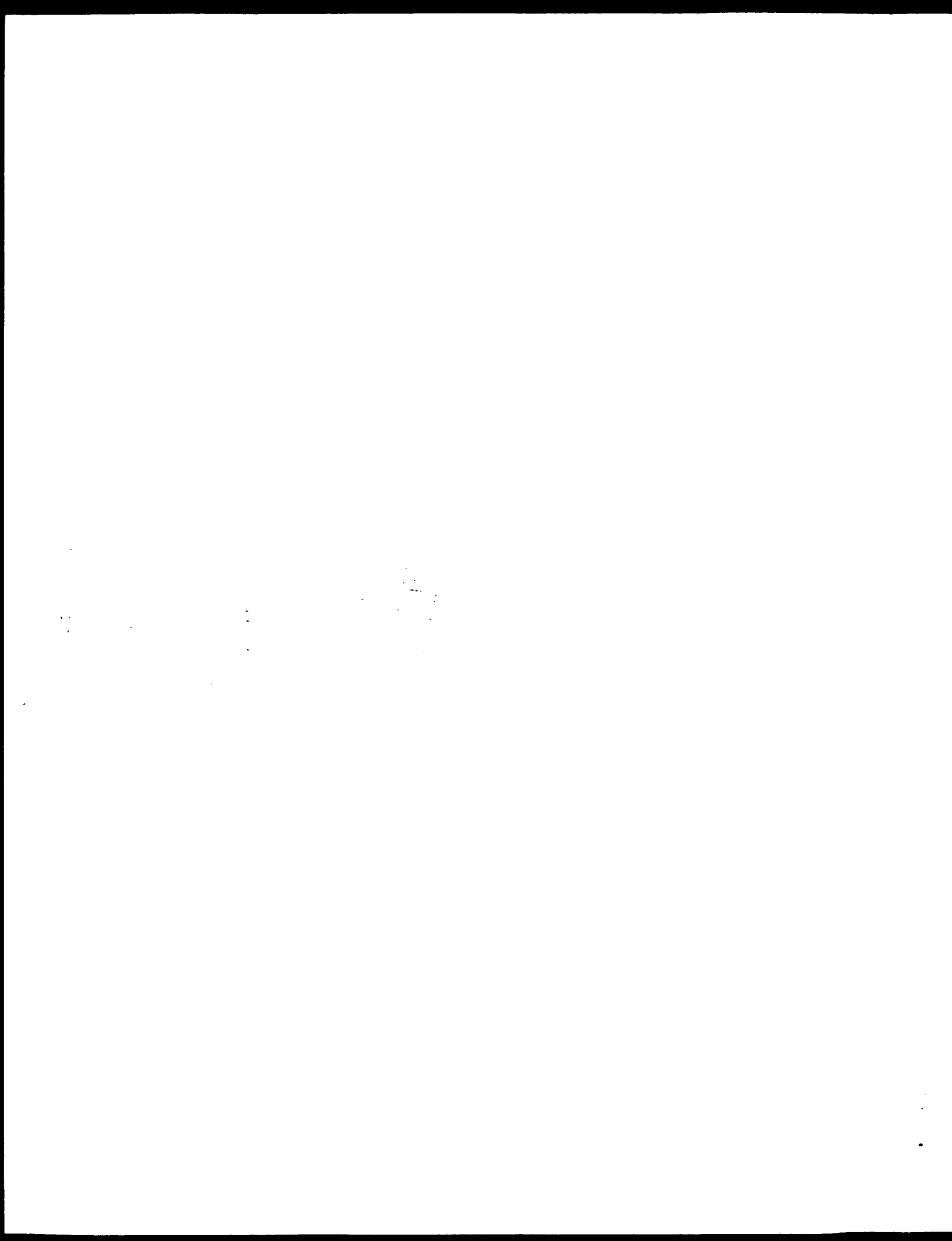
97 2960 CATTTAAATGTTAGTATGTTATTTATACAGATTTAAATTCCTGCTGCTGCTGCTGCTGCTGCT 3018
100 1 CATTTAAATGTTAGTATGTTATTTATACAGATTTAAATTCCTGCTGCTGCTGCTGCTGCTGCT 600

97 3019 CATTTAAATGTTAGTATGTTATTTATACAGATTTAAATTCCTGCTGCTGCTGCTGCTGCTGCT 3078
100 1 CATTTAAATGTTAGTATGTTATTTATACAGATTTAAATTCCTGCTGCTGCTGCTGCTGCTGCT 660

97 3079 ATAGTAAAGATG 3090
100 1 ATAGTAAAGATG 672

```

Search completed: January 18, 2003, 04:11:49
 Job time : 3148.5 secs



Genfore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:29:22 : Search time 2663.08 seconds
(without alignments)
1846% 506 Million cell updates/sec

Index: US-09-026-459A-34

Indexed score: 3266

Sequences: 1 GCGATGACGCGCGGCGAA.....AAATGAGATTAATGACT 3266

Sorted table:

IDENTITY: NUC

Gapop 10.0 , Gapext 1.0

Searches: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42408142

Minimum DB seq length: 0

Maximum DB seq length: 2000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: cm_estbacc*
- 2: cm_estbacc*
- 3: cm_estbacc*
- 4: cm_estbacc*
- 5: cm_estbacc*
- 6: cm_estbacc*
- 7: cm_estbacc*
- 8: cm_estbacc*
- 9: cm_estbacc*
- 10: cm_estbacc*
- 11: cm_estbacc*
- 12: cm_estbacc*
- 13: cm_estbacc*
- 14: cm_estbacc*
- 15: cm_estbacc*
- 16: cm_estbacc*
- 17: cm_estbacc*
- 18: cm_estbacc*
- 19: cm_estbacc*
- 20: cm_estbacc*
- 21: cm_estbacc*
- 22: cm_estbacc*
- 23: cm_estbacc*
- 24: cm_estbacc*
- 25: cm_estbacc*
- 26: cm_estbacc*
- 27: cm_estbacc*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	919.2	27.9	1624	BM450031	BM450031 AGENCOURT
2	909.4	27.8	1051	BM561083	BM561083 AGENCOURT
3	856	26.2	1981	BQ49221	BQ49221 AGENCOURT
4	844.8	25.9	881	BQ220275	BQ220275 AGENCOURT
5	786	24.1	1001	BM475603	BM475603 AGENCOURT
6	777.8	23.8	1364	AK011246	AK011246 Mus muscu

7	752.4	23.0	1122	13	BM545681	BM545681 AGENCOURT
8	725.4	22.2	727	9	AL597811	AL597811 DKFZ0136
9	720.6	22.1	820	14	BQ222227	BQ222227 AGENCOURT
10	699.6	21.4	823	12	BG610661	BG610661 602611848
11	698.2	21.4	792	9	AL563757	AL563757 AL563757
12	692.2	21.2	767	12	BQ570156	BQ570156 602791412
13	691.8	21.2	972	13	BM466194	BM466194 AGENCOURT
14	678	20.8	835	13	BM453724	BM453724 AGENCOURT
15	660	20.2	672	13	B1868459	B1868459 603492480
16	643.6	19.7	797	13	BE599278	BE599278 601064567
17	642.2	19.7	621	12	BG616219	BG616219 602644217
18	617	18.9	618	13	BM264009	BM264009 1044108.X
19	614	18.8	694	10	HE082846	HE082846 RC2 H1064
20	605.2	18.5	623	10	AW583141	AW583141 1a11109.Y
21	605	18.5	605	9	AL599633	AL599633 DKFZ014P
22	596.4	18.3	922	12	BQ353543	BQ353543 602464278
23	580	17.8	580	9	AL599105	AL599105 DKFZ0136
24	575.2	17.6	914	9	AA763411	AA763411 VW54404.Y
25	566.4	17.3	588	10	AA763485	AA763485 VW54401.Y
26	558	17.1	871	9	AA763485	AA763485 VW54401.Y
27	540.2	16.2	598	10	AW582967	AW582967 1a08608.Y
28	530	16.2	539	9	AL693700	AL693700 DKFZ0141
29	528	16.2	572	9	AL120941	AL120941 DKFZ0762D
30	516.8	15.8	597	12	BE931236	BE931236 MR0 H1015
31	513.2	15.7	594	10	AV715533	AV715533 AV715533
32	512.2	15.7	734	12	BF344857	BF344857 601790443
33	505.2	15.5	736	13	B1151568	B1151568 602916108
34	493.8	15.1	525	9	AA258255	AA258255 2159004.Y
35	486.4	14.9	584	12	BG149950	BG149950 0088-01.Y
36	484.4	14.8	493	10	AW502887	AW502887 0134P.HN0
37	481.4	14.7	484	10	HE168095	HE168095 0V3 H1051
38	480.2	14.7	659	13	B1821116	B1821116 603035591
39	476.6	14.6	2083	13	BM456728	BM456728 AGENCOURT
40	474.4	14.5	734	13	B1526682	B1526682 602927377
41	459.2	14.1	497	12	BE944784	BE944784 RC3.H1094
42	457	14.0	457	9	AL094215	AL094215 1004302.X
43	457	14.0	732	9	AL598766	AL598766 DKFZ0134
44	454.2	13.9	869	12	HE865749	HE865749 60167814
45	447	13.7	447	9	AT082179	AT082179 0A75604.X

ALIGNMENTS

RESULT 1
BM450031
LOCUS AGENCOURT_6392544 NIH_MGC_72 Homo sapiens cDNA clone IMAGE5728037
DEFINITION 1024 bp. mRNA linear EST 05-SEP-2002
5' mRNA sequence.
ACCESSION BM450031.1 GI:18499071
VERSION BM450031
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 1024)
Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
Enkaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
NHL-MGC http://www.ncbi.nih.gov/EST/18499071
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rstraus@nsl.nih.gov
Tissue Procurement: ATCC/DCT/DTP
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LAM1223 Row: 3 Column: 22
High quality sequence stop: 643.
location/Qualifiers
1. 1023


```

14 121 AATTATTAAGAAAATTCTCAATGAAACAAATTTTTCATATATGCTTTTATTGAGGTGGAGTCTT 180
15 1192 GAGGCTGCAAGGAAATATATATAGAGAGACATATCTAGAAATCTGATCTGGAATAGAT 1251
16 181 GAGGCTGCAAGGAAATATATATAGAGAGACATATCTAGAAATCTGATCTGGAATAGAT 240
17 1252 TTGGCTTTTGGATGATCTGAATGCTTAATTTTAAAGGCTTTGATTTTAAAGGCTG 1411
18 241 TTGGCTTTTGGATGATCTGAATGCTTAATTTTAAAGGCTTTGATTTTAAAGGCTG 400
19 1412 ATCTAAAGCTTTTAAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTG 1371
20 401 ATCTAAAGCTTTTAAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTG 360
21 1372 GCAATGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1431
22 461 GCAATGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 420
23 1442 GTTATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTT 1491
24 421 GTTATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTT 480
25 1492 GTTATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTT 1551
26 481 GTTATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTT 540
27 1552 TCTTAAAGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 1611
28 541 TCTTAAAGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 600
29 1612 GCAATGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1671
30 601 GCAATGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 660
31 1672 AAAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 1731
32 661 AAAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 720
33 1732 GCAATGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 1791
34 721 GCAATGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 780
35 1792 TATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAG 1851
36 781 TATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAG 840
37 1852 TATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAG 1910
38 841 TATTAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAG 900
39 1911 TCTTAAAGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 1962
40 901 TCTTAAAGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 953

```

RESULT 4
 HQ4.9221
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

1981 bp. mRNA. linear. EST: 24-MAY-2002
 AGENT: NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6015489
 5' mRNA sequence.
 HQ4.9221
 91:21179297
 ESI
 human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1081)
 NIH_MGC http://mgi.nhl.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rqualls@mail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (HLNI)
 DNA Sequencing by: Amersham Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/HLNI at:
 http://image.llnl.gov
 Plate: LLAM13211 row: k column: 01
 High quality sequence step: 742.
 Location/Qualifiers
 1..1081
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6015489"
 /cclone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage resistant)"
 /note="Organ: testis; Vector: pCMV-Sp-RP6; Site_1: Not 1;
 Site_2: Not 1; Cloned unidirectionally; cloned at primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by life technologies.
 Note: this is a NIH_MGC library."

BASE COUNT 337 a 190 c 183 q 371 t
 ORIGIN
 Query Match 26.2% Score 856; DB 14; Length 1081;
 Best local Similarity 99.1% Pred. No. 1,26-154;
 Matches 892; Conservative 0; Mismatches 5; Indels 3; Gaps 4;

QY 2368 GAAGCAGATGAGTAAATATCTCCACAGAGAGTCTAAATTTAGAGAGAAATGGAGAA 2427
 DB 1 GAAGCAGATGAGTAAATATCTCCACAGAGAGTCTAAATTTAGAGAGAAATGGAGAA 60
 QY 2428 ATGACCTCTACCTCAACATGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2487
 DB 61 ATGACCTCTACCTCAACATGAAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
 QY 2488 AACAG 2547
 DB 121 AACAG 180
 QY 2548 CATTCTCTCTACAGATGAGTGTATATCTTCCACAGTCTCTTATATATATATATAT 2607
 DB 181 CATTCTCTCTACAGATGAGTGTATATCTTCCACAGTCTCTTATATATATATATAT 240
 QY 2608 ATATCT 2667
 DB 241 ATATCT 400
 QY 2668 AAGCAGCTGGAATCTTACCTGATTTTATATATATATATATATATATATATATATAT 2727
 DB 301 AAGCAGCTGGAATCTTACCTGATTTTATATATATATATATATATATATATATATAT 460
 QY 2728 TGCCATTTAAAAAGTTGTAGCAGATTTGTTCTCTTCAAGTAAAAATGTGTGTATTA 2787
 DB 361 TGCCATTTAAAAAGTTGTAGCAGATTTGTTCTCTTCAAGTAAAAATGTGTGTATTA 420
 QY 2788 TGCCATTTAAAAAGTTGTAGCAGATTTGTTCTCTTCAAGTAAAAATGTGTGTATTA 2847
 DB 421 TGCCATTTAAAAAGTTGTAGCAGATTTGTTCTCTTCAAGTAAAAATGTGTGTATTA 480
 QY 2848 TGCCATTTAAAAAGTTGTAGCAGATTTGTTCTCTTCAAGTAAAAATGTGTGTATTA 2907
 DB 481 TGCCATTTAAAAAGTTGTAGCAGATTTGTTCTCTTCAAGTAAAAATGTGTGTATTA 540
 QY 2908 TTTTATTTAAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 2967
 DB 541 TTTTATTTAAATTAATTAAGGCTTTGATTTTAAAGGCTTTGATTTTAAAGGCTTTGAT 600
 QY 2968 TTTGATGCTGCTCCCAATCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3027
 DB 601 TTTGATGCTGCTCCCAATCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660


```

QY 607 ATAGCTTTTATGAAATTCCTGCTGACATGCTAATGACATGCTGACAGAGCTTGAATAAT 666
Db 421 ATAGCTTTTATGAAATTCCTGCTGACATGCTAATGACATGCTGACAGAGCTTGAATAAT 480
QY 667 CTTCCTAAAGGATGACAAATTTATCTTTAAATTAATGATCTAGATGTAAGATATTTT 726
Db 483 CTTCCTAAAGGATGACAAATTTATCTTTAAATTAATGATCTAGATGTAAGATATTTT 540
QY 727 TTGATCATGATATAAAGCTCTTCAGACTGATTTCTATATACAGCTTTTGAACACAGAGAAC 786
Db 541 TTGATCATGATATAAAGCTCTTCAGACTGATTTCTATATACAGCTTTTGAACACAGAGAAC 600
QY 787 CTACGAAAAAGTAAGCTTCATGACAGAGGATGATTAATCTCCACACACACACACACAGG 846
Db 601 CTACGAAAAAGTAAGCTTCATGACAGAGGATGATTAATCTCCACACACACACACACAGG 688
QY 847 ATGTGATGACACTATGACAAATTAATGATGATTTTAAATTCAGAGAGTGAACACAC 905
Db 661 ATGTGATGACACTATGACAAATTAATGATGATTTTAAATTCAGAGAGTGAACACAC 729
QY 906 TTCAAAAATGCTGATTTGCTATTTTAA 932
Db 721 TTCAAAAATGCTGATTTGCTATTTTAA 747

RESULT 10
LOCUS 602611848F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4737634 5'
DEFINITION mRNA sequence.
ACCESSION M62610661
VERSION M62610661.1 GI:1462032
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 823)
NIH-MGC: http://mgs.cni.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: etapbs-ramail.nih.gov
Tissue procurement: DCLID/OTF
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC1605 row: 6 column: 11
High quality sequence stop: 709.
Location/Qualifiers
1..823
/organism="Homo sapiens"
/ab.xref="Taxon:9606"
/clone="IMAGE:4737634"
/clone_1lib="NIH_MGC_60"
/zlib_host="DH10B (11 phage-resistant)"
/zlib_host="DH10B (11 phage-resistant)"
/zlib_host="DH10B (11 phage-resistant)"
/zlib_host="DH10B (11 phage-resistant)"
); Double stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCTATTATGCT-3' and 3' adaptor
sequence: 5'-ATCTACAGCGCGAGCGCGCGCATG-dT(30)BN-3'
(Where B = A, C, G and N = A, C, G or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: This is a NIH_MGC
library."
BASE COUNT 262 a 148 c 142 g 271 t

```

ORIGIN

```

Query Match 21.4%; Score 699.6; Db 12; Length 823;
Best Local Similarity 94.8%; Pred. No. 1.5e-124;
Matches 779; Conservative 0; Mismatches 34; Indels 9; Gaps 5;

QY 2394 AGCAGAGTCCAAATTTTCAGCAGAAAGCTGCGAAGTCTTCTACTCTGAAACAGAAATGCA 2453
Db 1 AGGAGAGTCCAAATTTTCAGCAGAAAGCTGCGAAGTCTTCTACTCTGAAACAGAAATGCA 60
QY 2454 AAAGCAGAAATGATGATAGCATGATAGCTCAAAATGAGAGAGAAATGAGATGCA 2513
Db 61 AAAGCAGAAATGATGATAGCATGATAGCTCAAAATGAGAGAGAAATGAGATGCA 120
QY 2514 GCACTTCTGAGAAATTTCTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2573
Db 121 GCACTTCTGAGAAATTTCTGATTAATGATGATGATGATGATGATGATGATGATGATGAT 180
QY 2574 TAACCTTCCAGGTTCTGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2633
Db 181 TAACCTTCCAGGTTCTGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 2634 AAATGTCAGATGCAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGATGAT 2693
Db 241 AAATGTCAGATGCAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGATGAT 300
QY 2694 TATTATACAGATGCAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGATGAT 2753
Db 301 TATTATACAGATGCAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGATGAT 360
QY 2754 GTTCTCTCTCCAAAGTAAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGAT 2813
Db 361 GTTCTCTCTCCAAAGTAAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGAT 420
QY 2814 GATCTCTCTCCAAAGTAAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGAT 2873
Db 421 GATCTCTCTCCAAAGTAAATTTGTTGGTGTATGATGATGATGATGATGATGATGATGAT 480
QY 2874 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2933
Db 481 TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540
QY 2934 TACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2993
Db 541 TACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 2994 TTATCTGTAAGTCTTCTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 3050
Db 601 TTATCTGTAAGTCTTCTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 660
QY 3051 ATTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3110
Db 661 ATTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 717
QY 3111 AAAATTTTCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 3170
Db 718 AAAATTTTCTTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 776
QY 3171 TACTATTTGAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 3232
Db 777 TACTATTTGAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAAT 816

```

RESULT 11

```

AL563757/c
LOCUS AL563757
DEFINITION AL563757
ACCESSION AL563757
VERSION AL563757.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```



```

10 61 TAAGCTTGGATTAATGCTCTCTCAAGAGAGAGAGATAGATAGATTCAGTTCTGTT 120
QY 2594 TATGGTCATATTAATAATGTCAGGCTCTTTTGGGATATAAAAAAGGACATCCCAATG 2653
DB 121 TATGGTCATATTAATAATGTCAGGCTCTTTTGGGATATAAAAAAGGACATCCCAATG 180
QY 2654 TTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 2713
DB 181 TTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 240
QY 2714 TTTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 2773
DB 241 TTTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 300
QY 2774 ATGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 2833
DB 301 ATGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 360
QY 2833 ATGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 2893
DB 361 ATGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 420
QY 2894 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 2953
DB 421 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 480
QY 2954 TTTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 3012
DB 481 TTTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 540
QY 3013 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 3072
DB 541 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 600
QY 3073 TTTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 3132
DB 601 TTTGGGTCATGTTAAGGCAATTAAGATGTTAGTCTGTTTATATACAGATGTAATA 659
QY 3133 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 3192
DB 659 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 718
QY 3193 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 3240
DB 719 AATTAATGATATATATATATATATATATATATATATATATATATATATATAT 767

RESULTS
LOCUS BM466194 972 bp mRNA linear EST 05-FEB-2002
DEFINITION ACN000001.04:6974 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5576599
5' mRNA sequence.
ACCESSION BM466194
VERSION BM466194.1 GI:18515246
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 972)
NIH_MGC http://mimc.nhlbi.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
unpublished (1999)
contact: Robert Strausberg, Ph.D.
Email: capus@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Amersham Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1A01249 Row: J Column: 12

```

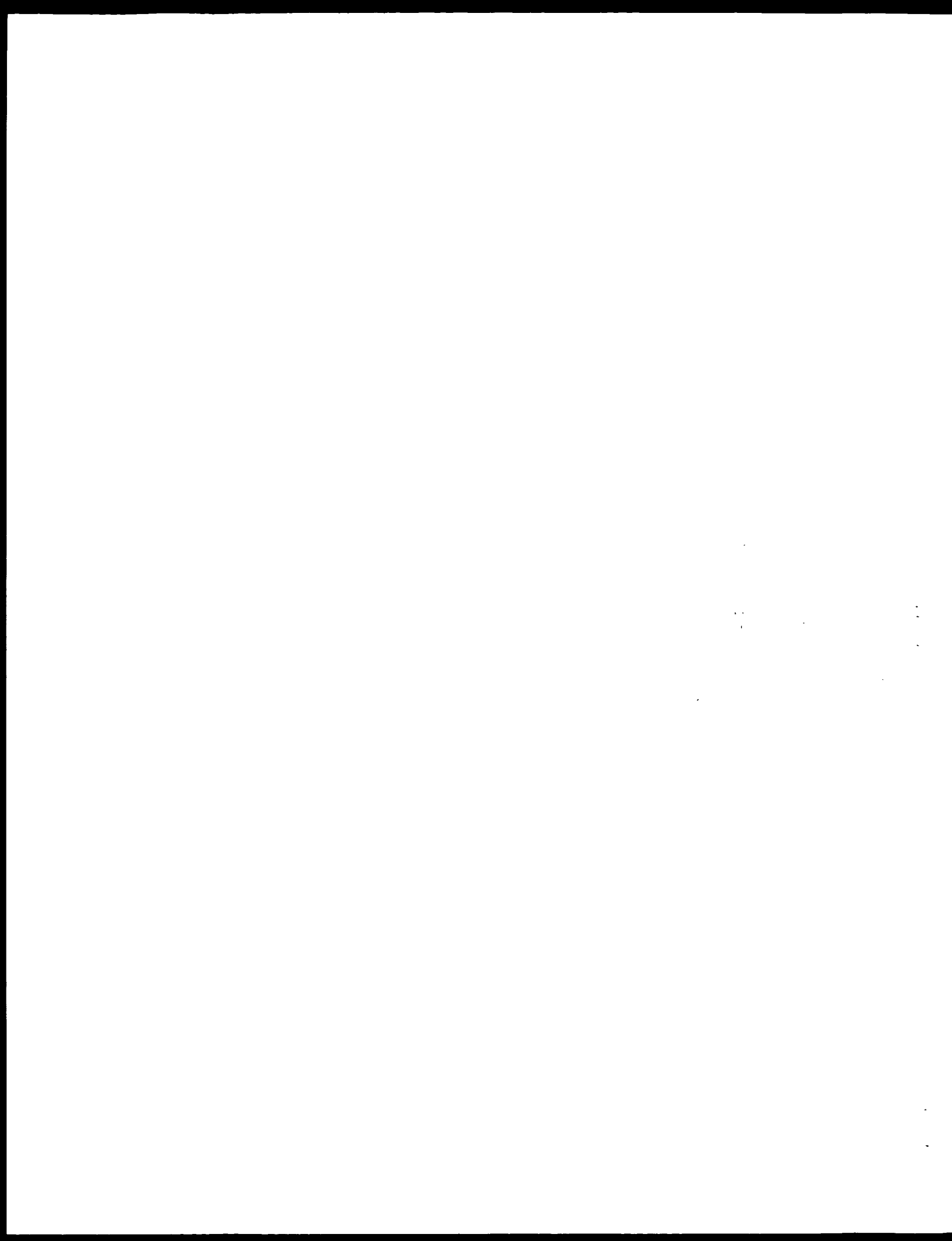
```

FEATURES             High quality sequence stop: 621.
                     Location/Qualifiers
                     1..972
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5576599"
                     /tissue_type="embryonal carcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Ordan; testis; Vector: pMW Shuttle; Site 1: Not 1;
                     Site 2: Salt; Cloned unidirectionally; original primed;
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies
                     Note: this is a NIH_MGC Library."
BASE COUNT          294 a 184 c 169 g 323 t 2 others
ORIGIN
Query Match      21.2% Score 691.8; DB 14; Length 972;
Best Local Similarity 96.7%; Pred. No. 4.7e-124;
Matches 728; Conservative 0; Mismatches 19; Indels 6; Gaps 2;
QY 2410 CAGTACAAACGGCAGAAACGACCTCTACATCAATATCAAAATCAAAATCAAAATCAAA 2469
DB 1 CATCAAAACGGCAGAAATGACTTCTACTGAAACAGAAATGAAAGAGAAAAATCAAT 60
QY 2470 GATACATGGATAGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2529
DB 61 CATACATGGATAGCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 2530 TGTGTACAGCTCTGGATTCATGCTCTCTCAAGATGAGATGATATATTTTGTGAGTTC 2589
DB 121 TGTGTACAGCTCTGGATTCATGCTCTCTCAAGATGAGATGATATATTTTGTGAGTTC 180
QY 2590 TGTGTACAGCTCTGGATTCATGCTCTCTCAAGATGAGATGATATATTTTGTGAGTTC 2649
DB 181 TGTGTACAGCTCTGGATTCATGCTCTCTCAAGATGAGATGATATATTTTGTGAGTTC 2709
QY 2650 ATTGTGTGATGATGCTTAAAGCAATTCGAAATGATGATGATGATGATGATGATGATG 2769
DB 241 ATTGTGTGATGATGCTTAAAGCAATTCGAAATGATGATGATGATGATGATGATGATG 300
QY 2710 AATATCTGTGTAAATCTGCTTAAATCTGCTTAAATCTGCTTAAATCTGCTTAAATCTG 2769
DB 301 AATATCTGTGTAAATCTGCTTAAATCTGCTTAAATCTGCTTAAATCTGCTTAAATCTG 360
QY 2770 TAAAAATGCTGTGCTTTTAAAGCAATTCGAAATGATGATGATGATGATGATGATGATG 2829
DB 361 TAAAAATGCTGTGCTTTTAAAGCAATTCGAAATGATGATGATGATGATGATGATGATG 420
QY 2830 AGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2889
DB 421 AGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
QY 2890 TATTAATTTATATGATATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 2949
DB 481 TATTAATTTATATGATATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 540
QY 2950 TCTATCTTCCAAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3009
DB 541 TCTATCTTCCAAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 3010 TGTAAAAATGATATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 3069
DB 601 TGTAAAAATGATATATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTT 660
QY 3070 ATTTTACTATTGGAAATCTGATATATGTTG-----TGTGTTTTTATATATATATATAT 3124
DB 661 ATTTTACTATTGGAAATCTGATATATGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 720
QY 3125 AATTAATTAAGAGCTTGAAGCAAGATTAAGCC 3156
DB 721 AATTAATTAAGAGCTTGAAGCAAGATTAAGCC 753

```


27 2192 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2251
DB 61 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 120
27 2252 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2311
DB 121 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 180
27 2312 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2371
DB 161 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 240
27 2372 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2431
DB 241 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 300
27 2432 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2491
DB 301 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 360
27 2492 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2551
DB 361 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 420
27 2552 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2611
DB 421 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 480
27 2612 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2671
DB 481 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 540
27 2672 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2730
DB 541 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 600
27 2731 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 2790
DB 601 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT 660
27 2791 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT
DB 661 TCAATCAAGATCAAGAACTTAGTATCAATATGTTGAATCATTCGGGAACTTTGAGAACT

Search completed: January 14, 2004, 04:07:55
Job time : 2894.96 secs




```

602 GCTGAAGGAGCAAGACGCTTTAAATGACGAGAAATAGCTTTGATGATTCAGAGATCA 721
QY 826 AspGlu 827
DB 122 CATAA 727

RESULTS 4
RM454724 835 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6402743 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528814
5' mRNA sequence
ACCESSION RM454724
VERSION RM454724.1 GI:184502764
KEYWORDS EST
SOURCE Human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 835)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstra@nhi.nih.gov
Tissue Procurement: ATCC/DCIB/HP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Applied Biosystems/TheraLabs
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: IAM12204 row: c column: 07
High quality sequence stop: 647.
Location/Qualifiers
1..835
/organism="Homo sapiens"
/db_xref="taxon:9606"
/zclone="IMAGE:5528814"
/zclone_1ib="NIH_MGC_72"
/zclone_type="melanotic melanoma"
/zlab_host="F010B (phage resistant)"
/znote="Skin; Vector: pCMV Sport6; Site: 1; Not1;
Site 2: SalI; Cloned unidirectionally. Primer: Gligo di.
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 254 a 210 c 137 g 243 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 8,056-110 Length: 835
Source: 1208-00 Matches: 244
Percent Similarity: 92.91% Conservatio: 5
Best Local Similarity: 91.94% Mismatches: 12
Query Matches: 26,878 Indels: 7
Dns: 13 Gaps: 5
US 09 026 459a 49 (1-831) x RM454724 (1-835)

QY 493 LeuGluAlaThrSerAlaPheGluThrGlnTyrProLeuYssSerProLeu 512
DB 8 GTCGACGAGAGGAGCAATGAAATGAGAAATGCGCTGCAATGCGCTGCAATGCGCTTAA 67
QY 513 PheAspLeuIleTyrGlnSerIyrAspArpIleGlyProThrAspIleLeuGluSerAla 522
DB 68 TTGATCTATTAAACAAATCAAGAGACGAGAGAGGACCACTGATCACTTGATCTGCT 127
QY 533 CysProLeuAsnGlnProLeuAspAsnIleValAlaAspMetIyrGlnSerPro 552
DB 128 TGTCTCTTAATCTCTCTCTGATGAAATATCACTGACAGATATGATCTCTCTCT 187
QY 554 ValAlaSerProIyrGlnTyrGlySerThrThrArpValIleAsnSerThrAlaAsnAlaGly 572
DB 188 GTAGATATGCAAGCAAAAGAAAGGCTTAACTACCGCTGCTAAATCTCTACTGCAATGACAG 247

```

```

QY 573 ThrGlnAlaThrSerAlaPheGluThrGlnTyrProLeuYssSerProLeuSerLeuSerLeu 592
DB 248 ACACAAAGCAACCTCAGCTTCGAGAGCCCAAGCAATGAGATGATCTCTTTTATG 407
QY 593 PheTyrIysIysValIyrArgLeuAlaTyrIleAsnIleAsnIleAsnIleAsnIleAsn 612
DB 308 TTTTATAAAAAAGGTATACGCTAGGCTATCTCGGCTAAATATACCTTCTGAAACCTT 467
QY 613 LeuSerGluIleHisProGluLeuGluHisIleIleTyrThrLeuPheGluHisThrLeuGln 632
DB 368 CTGCTCAGCAGCAGCAATAGACATATATCTGACCTTTCTCCACACACACCTCTGAG 427
QY 633 AsnGlnTyrGlnLeuMetArgAspArgHisIleAspIleIleMetIleCysSerMetIyr 652
DB 458 AATGAGTATCAACTCATGAGAGACAGGCAATTGACCAATATGATGCTCTCTATGAT 487
QY 653 GlyIleCysIysValIysAsnIleAspLeuIysPheIysIleIleValThrAlaIyrIys 672
DB 488 GCTATATGCAAACTCAACCAATATAGAGCTTAAATCTAAAACTATGTAATATATATAC 547
QY 673 AspLeuProHisAlaValGlnGluThrPheIysArgValIleValIleCysIleIleIyr 692
DB 548 GATCTCTCATGCTGCTTCCAGACACATTCACACGCTGCTTTGATCAAAAGCAATGAT 607
QY 693 AspSerIleIleValPheTyrAsnSerValPheMetGlnArgLeuIysThrAsnIleLeu 712
DB 608 GATCTATATATGATTTCTATAATCTGGTCTCTCATGCAAGACTGAAATATATATTTG 667
QY 713 GlyTyrAlaSerThrArgProPro---ThrLeuSerProIleProHisIleProArgSer 731
DB 668 CAGTATGCTTCCACAGCGCGCGCTATCCCTTGATCACTGCAATGATCACTATCTCTG 727
QY 741 rProTyrIys---PheProSerSerProLeu---ArgIlePro---GlyGlyAsn--- 746
DB 728 CCCCCTCAACACTTTCCTACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 787
QY 747 -IleTyrIleSerProLeuIys 753
DB 788 TATATTTTAAAGCCCTGAGAA 809

RESULTS 5
RM450031
LOCUS RM450031.1 GI:184599071
DEFINITION AGENCOURT_6493544 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528037
5' mRNA sequence
ACCESSION RM450031
VERSION RM450031.1 GI:184599071
KEYWORDS EST
SOURCE Human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1023)
AUTHORS NIH-MGC http://mgi.nhl.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rstra@nhi.nih.gov
Tissue Procurement: ATCC/DCIB/HP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Applied Biosystems/TheraLabs
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: IAM12204 row: d column: 22
High quality sequence stop: 643.
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/zclone="IMAGE:5528037"

```



```

146 424 ATATACATTTTGAACACACACAAACACACAAAAGTAGTAACCTTATGAAGAGGTGAAT 482
147 431 ValTleuProHisThrProValArgThrValMetAsnThrTleGluGluLeuMetMet 430
148 484 GTAAATCTCTACACACTAGTAGTAGAGCTGTATGAAACACTATGACAACTAATATGATG 542
149 431 TleuAsnSerAlaSerAspGluProSerGluAsnLeuTleSerTyrPheAsnAsnGly 450
150 543 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 602
151 541 ThrValAsnProLysLeuSerTleLeuLysArgValLysAspTleGlyTyrTlePheLys 470
152 604 ACATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 662
153 471 GluLysPheAlaTyrAlaValTleGlyTyrValGluTleGlySerGluArgTyrTyr 460
154 664 GAGAAATTTGTAAGAGTGAGGAGACAGAGTGTGTGAAATTAATTAATTAATTAATTAAT 722
155 491 LeuGlyValArgLeuTyrTyrArgValMetGluSerMetLeuLysSerGluGluGluArg 410
156 724 CTTCAGATGCTCTGATATACAGAGTAAAGGAAATCCATGCTTAAATCAAGAAACAGCA 782
157 411 LeuSerTleGluAsnPheSerLysLeuLeuAsnAsnTlePheHisMetSerLeuL 430
158 784 TATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 842
159 440 GATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 442
160 844 TAGGTGAGGTGTTTATGATTTGTAATGAGGACATATAGC 880

RESULT 7
AA764411/c
LOCUS
DEFINITION
IMAGE:1247598.5, Soares mammary gland NMI-MC Mus musculus cDNA clone
PROTEIN (HUMAN): gb:U66591. Mouse reticuloblastoma susceptibility
protein (MOUSE);, mRNA sequence.

ACCESSION
AA764411 GI:2814158
VERSION
AA764411.1
KEYWORDS
EST.
SOURCE
house mouse.
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 914)
Marra, M., Hillier, L., Allen, M., Howles, M., Dietrich, N., Dubouque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, R.,
Theising, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and
Waterson, R.
The WashU HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4434 Forest Park Parkway, Box 8001, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through HMI; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MOL:661286
Seq primer: 28413 rev2 kit from Amersham
High quality sequence stop: 422.
Location/Qualifiers
1..914
Zn-finger "Mus musculus"
Zn-finger "taxon:10090"
Zn-finger "IMAGE:1247598"
Zn-finger "Soares mammary_gland_NMI-MC"
Zn-finger "female (lactating)"
Zn-finger "mammary_gland"
FEATURES
source

```

```

/Tab_host="DB10R"
/Note="Vector: pT7TAP-Eae (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dG) primer. Double stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pT7TAP vector. Library is normalized. Library
was constructed by Benito Soares and M. Fatima Bonaldo."
BASE COUNT 257 a 178 c 217 g 261 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,186-94 Length: 914
Score: 1056.50 Matches: 244
Percent Similarity: 82.14% Conserved: 19
Host Local Similarity: 75.97% Mismatches: 53
Query Match: 23.50% Indels: 6
DB: 9 Gaps: 2
US-09-026-459A-49 (1-871) x AA764411 (1-914)
QY 279 PheLeuAspHisAspLysThrLeuGlnThrAspSerTleAspSerPheGlnThrGlnArg 298
DB 914 TTTTGGATGACGAAATTAACCTCGAGCTGT TCGTTTAGAGCTTTTAAATGAGAG 857
QY 299 ThrProArgLysSerAsnLeuAspGluGluValAsnValTleProProHisThrProVal 418
DB 856 ACCCCACGGAACACACCTGCTGT-GAAGTCGCAAACTGGGTACGTCGACACCTGGT 798
QY 319 ArgThrValMetAsnThrTleGlnGlnLeuMetMetTleLeuAsnSerAlaSerAspGln 438
DB 797 TGGCTGTTTGGACCTATCCCAACAAATTAATGTCATTTTAACTGGGCAAGACG 738
QY 339 ProSerGluAsnLeuTleSerTyrPheAsnAsnGlyThrValAsnProLysSerTle 358
DB 737 GAGTCAGAAATTCGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 679
QY 359 LeuLysArgValLysAspTleGlyTyrTlePheLysGluLysPheAlaTyrAlaValGly 478
DB 678 CTAAAGCGGCTT-AAAGACGCTGGCAGATGCTTAAACATGAACCTTCTA AAGCTGCTG 623
QY 379 GluGlyGlyValGluLysSerGluArgTyrTyrSerGluValAlaLeuArgTyrArg 498
DB 622 CCAGCTGCTGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 563
QY 399 ValMetGluSerMetLeuLysSerGluGluGluArgLeuSerTleGlnAsnProLys 418
DB 562 GTGATGGAATTCATGCTAGATAGAAAGAAAGAGTTTGTTCATTCAGAAATTAATGAAA 503
QY 419 LeuLeuAsnAspAsnTlePheHisMetSerLeuGluAlaLeuGluValMet 438
DB 502 CTCCTAAATGACAAATCTTCAATGCTTAACTGCTGCTGCTGCTGCTGCTGCTGCTG 443
QY 439 AlaThrTyrSerArgSerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPhePro 458
DB 443 GCTACCTATACACACAGACATATCCACCACTTCTGATTTGCAACACAGATTTGCTTCCTG 483
QY 459 TrpTleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValTleLysSerPhe 478
DB 382 TGGATCTGGAAGTACTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 423
QY 479 TleLysAlaGluGlyAsnLeuGluThrGluMetTleLysHisLeuGluArgTyrLysHis 498
DB 322 ATCAAGTCGGAAGCAACTTCACAGAGAAATGATAAACAATTAATTAATTAATTAATTAAT 263
QY 499 ArgTleMetGluSerLeuAlaTrpLeuSerAspSerPheLeuPheAspLeuTleLysGlu 518
DB 262 CGAATCAAGCAATCCCTTCATGCTTTCATGCTTTCATGCTTTCATGCTTTCATGCTTTCAT 203
QY 519 SerLysAspArgGluGlyProThrAspHisLeuGluSerAlaTyrProLeuAsnLeuPro 538
DB 203 TCTAAAGTATGAAAGAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 146

```



```

>>>FASTA:3M Homo sapiens
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Primates, Catartida, Haplorhina, Homo
1 (bases 1 to 588)
GCp: http://www.hudwig.org.br/GAPSPS
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL unpublished (1999)
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Hudwig Institute for Cancer Research
Rua Prof. Antonio Prudente, 169, 4 andar, 01506-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@hudwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.hudwig.org.br/GAPSPS)
181099-023-e10413-1999-10-18&t4-1)
Seq primer: puc 18 forward
High quality sequence stop: 559.
FEATURES
Source
1 588
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HT0181"
/dev_stage="Adult"
/zotc="Organ, head, neck, Vector, puc18, Site 1, Seal,
Site 2, Seal; A mini-library was made by cloning products
derived from APES-PCR (M.S. Letters Patent application
No 196,716 - Hudwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 187 a 139 c 100 g 162 t
ORIGIN
Alignment Scores:
Pred. No.: 4,2e 87 Length: 588
Score: 978.00 Matches: 191
Percent Similarity: 97.45% Conservative: 0
Best Local Similarity: 97.45% Mismatches: 4
Query Match: 21.75% Indels: 1
Gaps: 0
US-09-026-459a-49 (1-871) x AM36M234 (1-588)
>>>
606 TrpLeuSerArgSerProGluPheArgPheGluGlySerGlySerGlyPro 525
|||||
607 TGGTGTATATATATATATATATATATATATATATATATATATATATATAT 62
|||||
608 ThrAspHisThrGluSerProGluAspLeuProGluAsnAspHisThrAla 545
|||||
609 ACTATCACTGTAATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122
|||||
610 AlaAspMetTyrLeuSerProValArgSerProTyrGlySerThrThrArgVa 565
|||||
611 AGCACAATATATATATATATATATATATATATATATATATATATATATAT 182
|||||
612 LAspSerThrAlaAsnAlaThrGlnAlaThrSerAlaPheGlnThrGlnLysProLe 585
|||||
613 AAATCTACTGCAATATATATATATATATATATATATATATATATATATAT 242
|||||
614 uLysSerThrSerLeuSerGluThrTyrGlySerValTyrArgLeuAlaTyrLeuArgGle 605
|||||
615 GAAATATATATATATATATATATATATATATATATATATATATATATAT 402
|||||
616 uAsnThrLeuGlySerGluAlaLeuLeuSerGluHisProGlnLeuGlnHisThrThrTh 525
|||||
617 AAAATACATTCGGTGAACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 362
|||||
618 rLeuPheHisThrLeuAlaAsnThrGlnThrGlnThrLeuMetArgAspArgHisLeuAspGln 645
|||||

```

```

60 603 CTTTTCAGATATATATATATATATATATATATATATATATATATATAT 422
61 645 AlieMetMetCysSerMetTyrGlyGlyCysValGlyGlySerGlyGlyPhe 665
|||||
62 623 AATATGATGCTGCTATATATATATATATATATATATATATATATATAT 182
|||||
63 665 SttTleValThrAlaGlyTyrGlyAspLeuProGluHisAlaValGlnGlnThrPheLeuSerGluVa 685
|||||
64 683 AATATGTTTAAATATATATATATATATATATATATATATATATATATAT 542
|||||
65 685 lGcalleLysGluGlySerGlyTyrAspSerGlyGlyGlyGlyGlyGlyGly 700
|||||
66 643 TTTGATCATATATATATATATATATATATATATATATATATATATATAT 588
|||||
RESULT 14
A1646097 799 bp mRNA linear EST 29 APR-1999
x44097 xl Scores_mammary_gland_NLMG M33 macculus cDNA clone
IMAGE:1247677 3' similar to gb:M15400 RETINOBLASTOMA ASSOCIATED
PROTEIN (HUMAN); gb:M26391 Mouse retinoblastoma susceptibility
protein (MOUSE);, mRNA sequence.
ACCESSION A1646097
VERSION A1646097.1 GI:4724572
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Rodentia; Sciurognathi; Muridae; Mus;
1 (bases 1 to 709)
REFERENCE N-155AP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:561365
This clone was previously sequenced on the 5' end only, this now
data is from the 3' end
Possible reversed clone. PolyA not found
High quality sequence stop: 369.
FEATURES
Location/Qualifiers
1..709
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="IMAGEL1247677"
/score_lib "Scores_mammary_gland_NLMG"
/sex="female (lactating)"
/tissue_type="mammary gland"
/lab_host="pH10H"
/molecule="cDNA"
/vector="pUC19"
/notes="Vector: pUC19-pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified pUC19 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Honaldo."
BASE COUNT 211 a 150 c 159 g 185 t 5 others
ORIGIN
Alignment Scores:
Pred. No.: 7,73e-86 Length: 709
Score: 966.50 Matches: 187
Percent Similarity: 90.2% Conservative: 16
Best Local Similarity: 83.1% Mismatches: 21
Query Match: 23.5% Indels: 1
Gaps: 1
US-09-026-459a-49 (1-871) x A1646097 (1-709)
>>>
61 16 AlaAlaGluProProAlaProProProProProProProProProProPro 35

```


Search completed: January 19, 2003, 05:07:26
Job time : 1939.01 secs

bioWare version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

cdm protein nucleic search, using frame_plus_p2m model

Run on: January 17, 2003, 18:27:24 : Search time 1922:51 Seconds
(without alignments)
7320,587 Million cell updates/sec

Title: US-09-026-459A-47
Perfect score: 4302
Sequence: 1 MPKPKTKTAATAAAAAAP.....TRMKOKNNUSMDSNKEEK 869

Scoring table:
RGSUM62 Xgapop 10.0 : Xgapext 0.5
Ygapop 10.0 : Ygapext 0.5
Egapop 6.0 : Egapext 7.0
Dgapop 6.0 : Dgapext 7.0

Searched: 1615406 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32408132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Fast processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-s 5 -r 2 -l 1000 -p 20 -m 1000 -v 1000 -w 1000
DB-EST: OPTI-lastap SUFFIX-est -MINMATCH=0.1 -LOOP=0 -LEAP=0
LOCALISATION-START-1-ENCL-1 -MATRIX-biolum62 -TRANS-human40.edi -LIST-45
LOCALISATION-200 -THR_SCORE=100 -THR_MAX=100 -THR_MIN=0 -ALIGN-1% -MODE=LOCAL
OUTFILE-ptg NORM-ext HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
USER:US-09-026-459A-47 -P2M -P2M -P2M -P2M -P2M -P2M -P2M -P2M -P2M -P2M
NO XINDEX NO MMAP LAPPEPEPE NPT-SCOPES=0 WAIT -1000000 -DEVT-TIMEOUT=120
WAYN TIMEOUT=40 THREADS=1 XGAPOP=10 XGAPEXT=0.5 XGAPOP-6 -FGAPEXT-7
YGAPOP=10 YGAPEXT=0.5 DELOP=6 DELEXT=7

Database:

- 1: cm estbat *
- 2: cm esthum *
- 3: cm estint *
- 4: cm estint *
- 5: cm estint *
- 6: cm estint *
- 7: cm estint *
- 8: cm hrc *
- 9: qb est1 *
- 10: qb est1 *
- 11: qb hrc *
- 12: qb est1 *
- 13: qb est1 *
- 14: qb est1 *
- 15: cm estint *
- 16: cm estint *
- 17: qb qss *
- 18: cm qss hum *
- 19: cm qss hum *
- 20: cm qss hum *
- 21: cm qss hum *
- 22: cm qss hum *
- 23: cm qss hum *
- 24: cm qss hum *
- 25: cm qss other *
- 26: cm qss prot *
- 27: cm qss prot *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1575.5	45.0	1364	11	AK011246 Mus musculus
2	1551	34.5	1051	13	H0561083 AGING-001
3	1278.5	28.4	861	14	H0220275 AGING-001
4	1254	27.9	727	9	AL597811 IRT2p4140
5	1208	26.8	845	13	H0454724 AGING-001
6	1111	24.7	1024	13	H0450041 AGING-001
7	1056.5	23.5	914	9	AA763411 vW54a04.1
8	1037	23.0	605	9	AL599634 IRT2p4140
9	1034	22.9	792	9	AL563757 AL563757
10	1030	22.9	694	10	H080846 RC2-R1064
11	1016	22.6	623	10	AW084181 IRT11009.1
12	1013	22.5	871	9	AA763445 vW54a04.1
13	1004	22.3	922	12	H0254543 G02364278
14	996.5	22.1	1001	13	H0475603 AGING-001
15	978	21.7	588	10	AW084244 CMA-RT018
16	962	21.4	746	13	H1151568 G02916108
17	959	21.3	594	10	AV715533 AV715533
18	955	21.2	584	12	H0149050 H08601.1
19	947.5	21.0	734	12	H0144857 G01790483
20	935	20.8	539	9	AL592700 IRT2p4140
21	908	20.2	797	10	H0559278 G01063567
22	894	19.9	883	14	H0222227 AGING-001
23	890	19.8	1154	14	H0802736 AGING-001
24	882	19.6	572	9	AL1120941 IRT2p4140
25	879.5	19.5	734	13	H0526082 G02527277
26	859.5	19.3	845	12	H0244348 G0244348
27	858	19.1	592	12	H0231336 MRO-HT015
28	844	18.7	404	10	H0168095 QV3-HT051
29	833.5	18.5	645	13	H0490416 IRT2p4140
30	824	18.3	473	13	AW502887 IRT2p4140
31	818	18.2	832	9	AL466648 vW54a04.1
32	815.5	18.1	522	9	AA072789 G01790483
33	802	17.8	643	13	H0554937 G01790483
34	784	17.4	2083	13	H0456228 AGING-001
35	760	16.9	615	9	AA59836 vW54a04.1
36	755	16.8	709	9	AL466097 vW54a04.1
37	751	16.7	453	10	AW50514 IRT2p4140
38	703.5	15.6	472	9	AA955457 IRT2p4140
39	691	15.3	514	12	H0418684 IRT2p4140
40	684	15.2	448	12	H0227047 RC3-H0002
41	681.5	15.1	799	14	CH2304 CH2304 Leuk
42	676	15.0	545	10	H0592220 G02916108
43	667	14.8	651	13	H0696841 G03444967
44	659	14.6	480	10	AW951604 EST363674
45	633	14.1	425	10	BE285796 G01096110

ALIGNMENTS

RESULT 1

AK011246

LOCUS

DEFINITION

AK011246
Mus musculus 10 days embryo whole body cDNA, RIKEN full length
cDNA library, clone:260017M22;rat10blastoma 1, full insert
sequence.

ACCESSION

AK011246

VERSION

AK011246.1 G1:12847.243

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (strain:G78076) 10 days embryo cDNA to mRNA,
clone:lib:RIKEN full length enriched mouse cDNA library
clone:260017M22.

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1364 bp mRNA linear, RIKEN full length
cDNA library, clone:260017M22;rat10blastoma 1, full insert
sequence.

1. Carriani, P. and Pignatelli, B. V.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 JOURNAL
 PubMed
 10449646
 2.
 Carriani, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL
 MEDLINE
 PubMed
 20499374
 3.
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carriani, P.,
 Konno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, K., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujikawa, S., Inoue, K., Nagawa, Y., Iizawa, M., Ohara, F., Warahiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsushima, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (PISA) system for rapid discovery of
 sequencing pipeline with 384 multi-capillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 JOURNAL
 MEDLINE
 PubMed
 20540933
 4.
 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
 Aizawa, K., Hara, A., Fukui, T., Nishi, Y., Konno, H., Adachi, J., Fukuda, S.,
 Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, T.,
 Saito, T., Okazaki, Y., Goto, T., Bono, H., Kasukawa, T., Saito, R.,
 Kado, T., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
 Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
 Kusl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pasolunghi, G.,
 Quackenbush, J., Schriml, J., Staub, J., Suzuki, P., Tomita, M.,
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
 Baldarelli, R., Harsh, C., Blake, J., Boffelli, D., Hoshino, N.,
 Carriani, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
 Holtzman, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
 Marchionni, L., Mashima, J., Mazzarelli, J., Montauti, F., Nardone, P.,
 Kind, B., Kinsaid, M., Rodriguez, J., Sakamoto, N., Sasaki, H.,
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. P., Suzuki, H.,
 Toyazaki, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
 Wyshaw-Boriss, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S.,
 and Hayashizaki, Y.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409 (6821), 645-649 (2001)
 JOURNAL
 MEDLINE
 PubMed
 21085660
 5. (bases 1 to 1364)
 Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Haldar, K., Bono, H., Brownstein, M., Bult, C.,
 Carriani, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, K., Kasukawa, T.,
 Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
 Numazaki, P., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
 Saito, H., Saito, K., Sakai, K., Sakai, K., Sano, H., Sasaki, D.,
 Schriml, J., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
 Sorate, Y., Sasaki, H., Tagawa, M., Tagawa, A., Takahashi, F.,
 Tanaka, T., Tejima, T., Toyota, T., Yamamura, T., Yamanaka, T.,
 Yasushiji, A., Yoshida, K., Yoshino, M., Yuzugata, M. and
 Hayashizaki, Y.
 Direct Submission
 Submitted (10-Jul-2000) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN) Laboratory for Genome
 Exploration and Research group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22, Esato-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0445, Japan (E-mail: genome-test@sc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/, tel: 045-504-9222)

QY	745	11cvtYrIleScritProLeuTys 751
DB	788	TATAATTTTCAACCGCTGGAA 809
RESULT 6		
BM450031		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		
FEATURES		
Source		
BASE COUNT		
ORIGIN		
Alignment Scores:		
Prod. No.:		
Score:		
Percent Similarity:		
Best Local Similarity:		
Query Match:		
DB:		
US-09-026-459A-47 (1-869) x BM450031 (1-1023)		
QY	654	LysValLysAsnIleAspLeuLysIleValThrAlaTyLysAspLeuPro 673
DB	2	AAATGGAAGATATAGACCTTAAATTCATAATCATTGTAACAGCATCAAGATCTCT 61
QY	674	HisAlaValGlnGluThrPheLysArgValIleCysGluCysGluThrAspSerIle 693
DB	62	CATCTCTTCAGACACATTCATAACGCTTTTTCATCAAGAGAGAGATGATGATTTAT 121
QY	694	IleValPheTyAsnSerValPheMetIleArgLeuLysThrAsnIleLeuGlnIleVal 713
DB	122	ATAGTATCTATAACTCTCTCTTCTATCATACAGATCAAAACAAATATTTCCAGATGCT 181
QY	714	SerThrArgProProThrLeuSerProIleProHisIleProArgSerProTyLysPhe 743
DB	182	TCCACAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241
QY	734	ProSerSerProLeuArgIleProGlyCysIleAsnIleTyLeuSerIleProGlySer 753


```

57  LysGlySerThrThrAlaValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 576
85  AGAAGTCTACACAGCTGTAATCTGTCGAATATGACAGACACAAAGCAGCCACAGCC 26
97  PheGlnThrGlnLysProLeuLys 584
11  LysGlnThrGlnLysProLeuLys 584
25  TCCATATGCAAGAGCAATGAAA 2

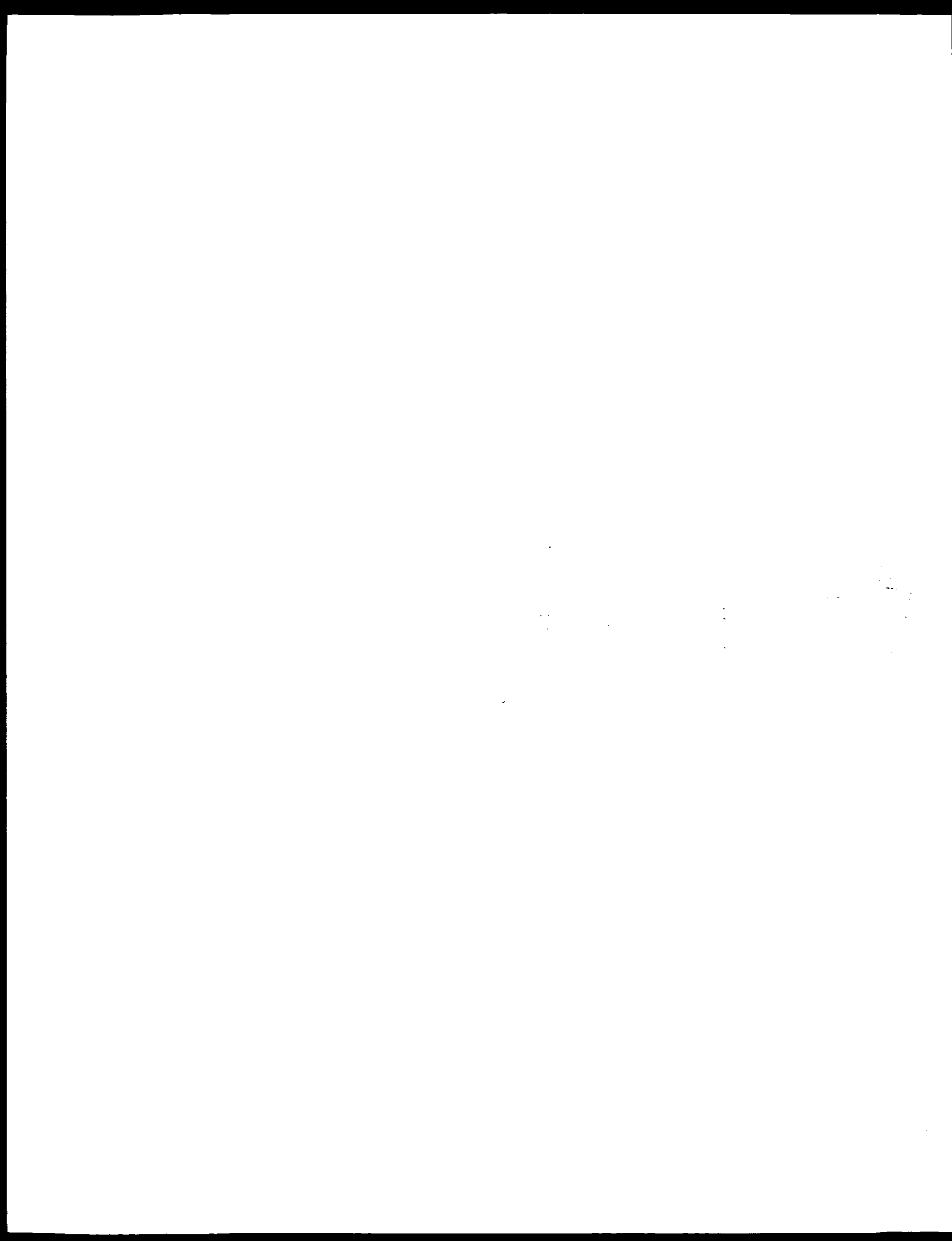
RESULT 8
AL599633 605 bp mRNA linear EST 14-AUG-2001
CYFIP13p092_1_13 (Cyfip13p092) Homo sapiens cDNA clone
DKE2p313p092 5', mRNA sequence.
AL599633
AL599633.1 GI:15162921
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo
1 (bases 1 to 605)
Koehler, K., Royer, A., Mewes, W., Weil, B. and Wiemann, S.
EST (Koehler, K., Royer, A., Mewes, W., Weil, B. and Wiemann, S.)
Unpublished (1999)
Contact: Koehler K
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available
This clone (DKE2p313p092) is available at the K2P6 in Berlin
Please contact the RPP: Ressourcenzentrum, Heuberweg 5, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rpp.de
Location/Qualifiers
1..605
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKE2p313p092"
/clone_lib="t13 (cyfip13p092, hlec2)"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pTrp1EX2; Site1, Site2, Site3"
cDNA-collection"
BASE COUNT 228 a 92 c 106 q 179 t
ORIGIN
1..605
ATTGATGCTTTTATGAATCTCTTGACCTGTGAACATCTAATGAGCTTCCAGAGGT 120
256
276
121

```

```

276  LeuPheLeuAspLysThrLeuGlnThrAspSerLleAspSerPheCl ThrGln 295
181  TTAATTTGGATCATGATAAAACTCTTCACACTGATCTATAGACAGTTTCAACACAG 240
296  ArgThrProArgLysSerAsnLeuAspGluGluValAsnValLleProProHisThrPro 315
241  ACAAACACACGAAAGGTAACCTTGATGAAGAGGTGAATTAATTCCTCCACACACTCA 300
316  ValArgThrValMetAsnThrLleGlnGlnLeuMetMetLleLeuAsnSerAlaSerAsp 335
301  GTTACGACTCTTATGAAACACTATCCCAACAAATTAATCATCATTTTAAATTCACCAACTCA 360
336  GlnProSerGlnAsnLleLleSerTyTheAsnAsnGlnLleValAsnProGlnLleSer 355
361  CAACCTTCACAAAATCTGATTTCCCTATTTTAAACAACATCCACAGTGAATCCAAAAGAAAG 420
356  LleLeuLysArgValLysAspLleClyTyrLlePheLysGlnLysPheAlaLysAlaVal 375
421  ATACTGAAAGAGTGAAGGATATAGCATACATCTTTAAAGAGAAATTTGCTAAGACTG 480
376  GlyGlnClyTyrValProLleClySerGlnAsnTyrLleSerGlnValArgLleLys 395
481  CGACAGCTGTCGTCGTAATTCGATACACATATCAAACTTCGAGTTCCGCTGCTATTA 540
396  ArgValMetGlnSerMetLeuLysSerGlnGluGlnLeuSerLleGlnAsnLysSer 415
541  CGAGTAATGGAATCCATGCTTAAATCAGAGAGAGAGATTAATCCATTCAAAATTTAG 600
416  Lys 416
601  AAA 603
RESULT 9
AL563757/c
LOCUS
DEFINITION
AL563757 AL563757.1 Homo sapiens cDNA clone cs00pp07YA05.3
prime, mRNA sequence.
AL563757
AL563757.1 GI:12913464
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo
1 (bases 1 to 792)
Li, W.R., Gruber, C., Jessee, J. and Pelayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
Location/Qualifiers
1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS00p07YA05"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Oligo: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
cloned, double stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen, 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: fliang@lifestech.com URL:
http://Fulllength.invitrogen.com"
BASE COUNT 245 a 148 c 117 q 255 t 26 others
ORIGIN

```

GenCore version 5.1.4
Copyright (c) 1993 - 2003 CompuGen Ltd.

EM nucleic acid sequence search, using SW model

Run on: January 16, 2003, 15:20:22 : Search time 201.21 seconds
(without alignments)
18469.506 Million cell updates/sec

Filter: US 09 026 459A 46

Perfect score: 3377
Sequence: 1 CGGCGTCAAGCGGCAAAAC.....AAATGAGGATTAATGATGACT 3377

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searches: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32408192

Minimum hit seq length: 0
Maximum hit seq length: 20000000

Post processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database: EST

- 1: em_ostba1 *
- 2: em_osthum *
- 3: em_ostlin *
- 4: em_ostma *
- 5: em_ostov *
- 6: em_ostpl *
- 7: em_ostro *
- 8: em_hic *
- 9: qb_est1 *
- 10: qb_est2 *
- 11: qb_hic *
- 12: qb_est3 *
- 13: qb_est4 *
- 14: qb_est5 *
- 15: em_ostum *
- 16: em_ostom *
- 17: qb_oss *
- 18: em_oss_hum *
- 19: em_oss_inv *
- 20: em_oss_pln *
- 21: em_oss_vit *
- 22: em_oss_tan *
- 23: em_oss_mam *
- 24: em_oss_mus *
- 25: em_oss_het *
- 26: em_oss_pro *
- 27: em_oss_rod *

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.2	27.0	1023	13	BM450031 AGENCOURT
2	907.4	26.9	1051	13	BM561083 AGENCOURT
3	896	26.4	1081	14	BQ439221 AGENCOURT
4	742.4	22.3	1122	13	BM545881 AGENCOURT
5	740.8	21.9	981	14	BQ220275 AGENCOURT
6	725.4	21.5	727	9	AL597811 DKF2p3130

7	699.6	20.7	923	12	BQ610961
8	692.2	20.5	767	12	BQ570456
9	691.8	20.5	972	13	BM466194 AGENCOURT
10	691.2	20.5	792	9	AL563757
11	678	20.1	835	13	BM453724
12	660	19.5	672	13	BH868459
13	646	19.1	1364	11	AK011246
14	643.6	19.1	797	10	BQ539278
15	642.2	19.0	681	12	BQ616219
16	617	18.3	618	13	BM264009
17	614	18.2	634	10	BE082846
18	605.2	17.9	623	10	AW583181
19	605	17.9	605	9	AL599633
20	596.4	17.7	922	12	BQ253743
21	580	17.2	580	9	AL599105
22	575.2	17.0	914	9	AA763411
23	566.4	16.8	588	10	AA768234
24	558	16.5	871	9	AA763485
25	530.2	15.7	539	9	AL692790
26	530	15.7	539	9	AL692790
27	528	15.6	572	5	AL220941
28	516.8	15.3	597	12	BF931246
29	513.2	15.2	594	10	AV715533
30	512.2	15.2	734	12	BF344857
31	508.6	15.1	1154	14	BH002746
32	505.2	15.0	736	13	BH151568
33	493.8	14.6	525	9	AA258255
34	486.4	14.4	584	12	BG149050
35	484.4	14.3	493	10	AW502887
36	484.2	14.3	1001	13	BM475603
37	481.4	14.3	484	10	BE168095
38	480.2	14.2	659	13	BH211116
39	476.6	14.1	2083	13	BM456728
40	474.4	14.0	734	13	BH526692
41	465.6	13.8	880	14	BQ222227
42	459.2	13.6	497	12	BQ934784
43	457	13.5	457	9	AL093215
44	457	13.5	732	9	AL598766
45	455.4	13.5	643	13	BH553637

ALIGNMENTS

RESULT 1
BM450031
LOCUS AGENCOURT_6493544 NIH_MGC_72 Homo Sapiens cDNA clone IMAGE:5528037
DEFINITION 1023 bp mRNA linear EST 05-FEB-2002
5' mRNA sequence
ACCESSION BM450031 GI:184599071
VERSION BM450031.1
KEYWORDS EST
SOURCE human
ORGANISM Homo Sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1023)
AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rapbs@mail.nih.gov
Tissue Procurement: ATCC/DCTL/DBP

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L2AM12254 row 3 column 22
High quality sequence stop: 643
Location/Qualifiers
source 1, 1023

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:552837"
/clone_lib="NIH_MGC_72"
/tissue_type="melanotic melanoma"
/label="lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-Sport6; Site: 1; Not1; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
BASE COUNT      344 a      211 c      197 g      281 t
ORIGIN

QY 1965 CAAAGTGAAGAAATATACAGCTTAAATTCAAATCAATGATGTAAGATACAGAGAGAGCTCC 2024
DB 1 CAAAGTGAAGAAATATACAGCTTAAATTCAAATCAATGATGTAAGATACAGAGAGAGCTCC 60
QY 2025 TCATGCTGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2084
DB 61 TCATGCTGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 2085 TATAGTATTCATATACCTGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2144
DB 121 TATAGTATTCATATACCTGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 2145 TTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2204
DB 181 TTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 2205 TCCATATACCTGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2264
DB 241 TCCATATACCTGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 2265 ATATATAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2324
DB 301 ATATATAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 2325 AGTATGAATTCGATCCTGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2384
DB 361 AGTATGAATTCGATCCTGCTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 2385 ATGTAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2444
DB 421 ATGTAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 2445 GAAAAAATAGGCTTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2504
DB 481 GAAAAAATAGGCTTCATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 2505 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2564
DB 541 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 2565 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2624
DB 601 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 2625 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2684
DB 661 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 2685 TAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2744
DB 721 TAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 2745 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2802
DB 781 AAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840

```

Query Match 27.0%; Score 910.2; DB 13; Length 1023;
Best Local Similarity 95.2%; Pred. No. 6.7e 168;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

RESULT 2
BM561083
DEFINITION 1051 bp mRNA sequence.
ACCESSION BM561083
VERSION 1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1051)
AUTHORS NIH-MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgs@nsl.nhl.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library: Arrayed by The I.M.A.G.E. Consortium (LMHL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNI at:
http://image.llni.gov
Plate: LLML2263 row: m column: 18
High quality sequence step: 737.

FEATURES
source
1..1051
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5550905"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: eye; Vector: pCMV-Sport6; Site: 1; Not1;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 330 a 221 c 197 g 299 t 4 others
ORIGIN

Query Match 26.9%; Score 909.4; DB 13; Length 1051;
Best Local Similarity 98.4%; Pred. No. 9.6e 168;
Matches 948; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1123 AAAATTCGCTGCTTCATGAG 1182
DB 1 AAAATTCGCTGCTTCATGAG 60
QY 1183 TCGATATTCAG 1242
DB 61 TCGATATTCAG 120
QY 1243 AATTTTGAAG 1302
DB 1 AATTTTGAAG 60

```

1421 AATTATGAAATTTCTGAAATGAAGAAATTTTTCATATATCTTTATTGGGAGGGCTCTT 180
QY 1403 CAGGTCGCTGAAATGCGACATATAGACAAAGTACATGACAAATCTGCAATTCGCAACAGAT 1462
Dd 1401 CAGGTCGCTGAAATGCGACATATAGACAAAGTACATGACAAATCTGCAATTCGCAACAGAT 240
QY 1463 TTGCTTTTGGATGATCTCTGATGCTTTAATTTTAAATGCTTTGATTTTACAAAGTG 1422.
Dd 241 TTGCTTTTGGATGATCTCTGATGCTTTAATTTTAAATGCTTTGATTTTACAAAGTG 300
QY 1423 ATGCAAACTGTTTAACTGCAAGAGTAACTTGCACAGACAAATGCAATTTACAA 1482
Dd 401 ATGCAAACTGTTTAACTGCAAGAGTAACTTGCACAGACAAATGCAATTTACAA 360
QY 1483 CTAATGTCGAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT 1542
Dd 661 CTAATGTCGAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT 420
QY 1543 CTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1602
Dd 421 CTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 1603 CTTAAATGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT 1662
Dd 481 CTTAAATGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAAT 540
QY 1663 TCTCAAGCAAAAGGCTCAACTACCGCTGCTAAATTTCTACTGCAATATGCAATATGCAAT 1722
Dd 541 TCTCAAGCAAAAGGCTCAACTACCGCTGCTAAATTTCTACTGCAATATGCAATATGCAAT 600
QY 1723 GAACTGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 1782
Dd 601 GAACTGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 660
QY 1783 AAAAAGGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 1842
Dd 661 AAAAAGGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 720
QY 1843 GAGTACGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 1902
Dd 721 GAGTACGCTGCTGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATG 780
QY 1903 TATGAAATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCA 1962
Dd 781 TATGAAATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCA 840
QY 1963 TATGAAATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCA 2021
Dd 841 TATGAAATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCA 900
QY 2022 TATGAAATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCA 2073
Dd 901 TATGAAATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCAATATGCA 953

RESULT 4
ACCESSION AC000001 7765662 NIH MGC_92 Homo sapiens cDNA clone IMAGE:5015480
VERSION B04.9221
KEYWORDS B04.9221.1 31.21179.27
SOURCE EST.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 1081)
AUTHORS NIH MGC; http://mgi.nhl.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: edupas@mail.nih.gov

```

```

FEATURES             Source
     Location/Qualifiers
     ..1..1081
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /clone_lib="IMAGE:6015480"
     /clone_lib="NIH MGC_92"
     /tissue_type="embryonal carcinoma, cell line"
     /lab_host="DH10B (blue-resistant)"
     /note="Organism: testis; pCMV Sport6; Site: 1; Note:
     Site_2: Salt; Cloned unidirectionally; oligo dT primed.
     Average insert size 2.5 kb. Library enriched for
     full-length clones and constructed by life technologies.
     Note: this is a NIH MGC library."
     BASE COUNT      337 a      190 c      183 g      371 t
     ORIGIN
     Query Match      25.3%   Score 856;   DB 14;   Length 1081;
     Best local Similarity 99.1%   P-adj. No. 2.7e-157;
     Matches 892;   Conservative 0;   Mismatches 5;   Indels 3;   Gaps 3;
QY 2479 GAAATGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2538
Dd 1 GAAATGATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 60
QY 2539 ATGACTTCTACTGCAACACCAATGCAAAAGTACAAATGCAATGCAATGCAATGCAATGCA 2598
Dd 61 ATGACTTCTACTGCAACACCAATGCAAAAGTACAAATGCAATGCAATGCAATGCAATGCA 120
QY 2599 AACAAAGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2658
Dd 121 AACAAAGCAAGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 180
QY 2659 CATTGCTCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 2718
Dd 181 CATTGCTCTGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 240
QY 2719 ATATCTTCAGCTCTTTTGGCAATATGCAATGCAATGCAATGCAATGCAATGCAATGCA 2778
Dd 241 ATATCTTCAGCTCTTTTGGCAATATGCAATGCAATGCAATGCAATGCAATGCAATGCA 300
QY 2779 AAGCACTTGAATGCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCT 2838
Dd 301 AAGCACTTGAATGCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCT 360
QY 2839 TGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2898
Dd 361 TGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
QY 2899 TGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2958
Dd 421 TGGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 2959 TGGCTCTTTTGGTACATATAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCT 3018
Dd 481 TGGCTCTTTTGGTACATATAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCT 340
QY 3019 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3078
Dd 541 TTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
QY 3079 TTTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3138
Dd 601 TTTCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 660

```

Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL4211 row: k column: 01
High quality sequence stop: 742.

Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

BASE COUNT 294 a 184 c 169 g 323 t 2 others

ORIGIN

```

Query Match      20.5%  Score 591.8;  Db 13;  Length 972;
Best local Similarity 96.7%  Pred. No. 3.6e-125;
Matches 726;  Conservative 0;  Mismatches 19;  Indels 6;  Gaps 2;

QY 2521 CAGAGAAATGGGAGAAATATTTATTCGAAACAGGAATGAAAAAGAGAAATATAT 2580
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2581 GATAGATAGGATACCTTAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
DB 1 GATAGATAGGATACCTTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 2641 TGTGTACAGCTGTGGATTTATTTCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2701 TGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2760 TGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2761 ATGTTATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2761 ATGTTATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2821 ATGTTATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 400
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2821 AAAAGCTGTGTAAATCTGCAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
DB 1 1111111111111111111111111111111111111111111111111111111
QY 301 AAAAGCTGTGTAAATCTGCAATTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
DB 1 1111111111111111111111111111111111111111111111111111111
QY 2881 TAAATATCTGTGTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
DB 1 1111111111111111111111111111111111111111111111111111111
QY 361 TAAATATCTGTGTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 1 1111111111111111111111111111111111111111111111111111111
QY 3941 AGGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
DB 1 1111111111111111111111111111111111111111111111111111111
QY 421 AGGCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 1 1111111111111111111111111111111111111111111111111111111
QY 6001 TATTAATTTATAGTATATTTTAAATTTATAGATAGAGAGAGAGAGAGAGAGAGAG 3060
DB 1 1111111111111111111111111111111111111111111111111111111
QY 481 TATTAATTTATAGTATATTTTAAATTTATAGATAGAGAGAGAGAGAGAGAGAGAG 540
DB 1 1111111111111111111111111111111111111111111111111111111
QY 6061 TATTAATTTATAGTATATTTTAAATTTATAGATAGAGAGAGAGAGAGAGAGAGAG 3120
DB 1 1111111111111111111111111111111111111111111111111111111
QY 541 TCTATCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 1 1111111111111111111111111111111111111111111111111111111
QY 4121 TGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
DB 1 1111111111111111111111111111111111111111111111111111111
QY 601 TGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 1 1111111111111111111111111111111111111111111111111111111
QY 4181 ATTTTATTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3235
DB 1 1111111111111111111111111111111111111111111111111111111
QY 461 ATTTTATTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 1 1111111111111111111111111111111111111111111111111111111
QY 4236 AATTAATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3267
DB 1 1111111111111111111111111111111111111111111111111111111
QY 721 ATTAAGATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353
DB 1 1111111111111111111111111111111111111111111111111111111

```

PEP811 10

AL564757

LOCUS

DEFINITION

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

AL564757

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates, Catarrhini, Hominidae; Homo.

REFERENCE 1 (bases 1 to 792)

AUTHORS Li, W. H., Gruber, C., Jessier, J., and Pelay, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage

at 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

Location/Qualifiers

1..792

source

/db_xref="taxon:9606"

/clone_lib="LTL_NFL001_NHC4"

/sex="male"

/issue_type="neuroblastoma cells"

/lab_host="PH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with NotI and

cloned into the NotI and EcoRV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact: Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive, Rockville, Maryland 20850, USA Fax: (301) 361-610

8371 Email: fliang@lifetech.com URL:

http://fulllength.invitrogen.com"

BASE COUNT 246 a 148 c 117 g 255 t 26 others

ORIGIN

Query Match 20.5% Score 601.2; Db 2; Length 792;

Best local Similarity 93.6% Pred. No. 4.8e-125;

Matches 715; Conservative 23; Mismatches 23; Indels 3; Gaps 2;

QY 553 GTGTTAATATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 512

DB 785 GCTTTATATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 726

QY 613 CGGATACAG 572

DB 725 CGGATACAG 666

QY 673 CAATGTAATATAGATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 732

DB 665 CAATGTAATATAGATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 606

QY 733 TATTTGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 792

DB 605 TATTTGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 546

QY 793 GAG 852

DB 545 GAG 486

QY 853 ACCTTTGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 912

DB 485 ACCTTTGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 426

QY 913 CTTCATGAG 972

DB 425 CTTCATGAG 366

QY 973 ATGAG 1032

DB 365 ATGAG 306

QY 1033 TCTTTGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 1092

DB 305 TCTTTGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTATATTAATGAGTAT 246


```

/lab_host="HOLIOH (phage-resistant)"
/notes="Organ: Liver; Vector: pCMV SVOR16; Site1: Not1;
Site2: SalI; Cloned unidirectionally; oligo dT primed.
Average insert size 1.7 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MRC library"
NAME: us09 222 a 134 g 186 t
BUILD

Query: Match 19.5% Score 660, DB 13; Length 672;
Best local similarity 99.9%; Pred. No. 6,1e-119;
Matches 671; Conservative 0; Mismatches 0; Indels 1; Gaps 1.

CY 243 ATATTTCAGGCTGAGAGAGGTCATATAAAATTTTCAGAGGCTGCTCCACACCAACAAAA 2302
EB 1 ATATTTCAGGCTGAGAGAGGTCATATAAAATTTTCAGAGGCTGCTCCACACCAACAAAA 60

CY 2403 TCAGTCCAAAGCAACAGAACTTACTATTAATTTGCTCAATCAATGCGGACATCTGAGAAAT 2352
EB 61 TCAGTCCAAAGCAACAGAACTTACTATTAATTTGCTCAATCAATGCGGACATCTGAGAAAT 120

CY 2463 TTTTAAATAATAATTCATATGTATATATATATATATATATATATATATATATATATATAT 2422
EB 1 TTTTAAATAATAATTCATATGTATATATATATATATATATATATATATATATATATATAT 180

CY 2423 CAAATCAATGCTGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 2482
EB 1 CAAATCAATGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 240

CY 2483 CAGATGAGAGAAACATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2542
EB 241 CAGATGAGAGAAACATCTCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300

CY 2543 GTTCTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA 2602
EB 401 GTTCTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA 360

CY 2603 AGCAATCAAAATTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2662
EB 461 AGCAATCAAAATTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

CY 2663 GCTCTCTACAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2722
EB 421 GCTCTCTACAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

CY 2723 GTTCTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA 2782
EB 481 GTTCTATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAAATTCGAA 540

CY 2783 CACTTCAAAATTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2843
EB 541 CACTTCAAAATTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

CY 2842 CATTTAAAAATTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2901
EB 601 CATTTAAAAATTCAGATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660

CY 2902 ATATTAATG 2913
EB 661 ATATTAATG 672

RESULT 14
AK011246
DEFINITION Mus musculus 10 days embryo whole body cDNA, RIKEN full length
enriched library, clone:260067M22-ref10ab2astoma 1, full insert
sequence.
ACCESSION AK011246
VERSION AK011246.1 GI:12847243
KEYWORDS HIC; CAP trapper.
SOURCE Mus musculus (strain:057H6/6J) 10 days embryo cDNA to mRNA,
clone_lib:RIKEN full-length enriched mouse cDNA library

```

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

clone:260067M22.

Mus musculus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10549636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sudahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1517-1530 (2000)

20499374

11042159

3 Shibata, Y., Itoh, M., Aizawa, K., Nishikawa, S., Sasaki, K., Ohtsuki, F.,

Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,

Sumi, K., Ishii, Y., Rakumura, S., Hama, M., Nishine, T., Harada, A.,

Yamamoto, P., Marumoto, H., Sakaguchi, S., Ikegami, T., Kachiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Warahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Inaka, T., Matsuura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Itoh, Y., Kita, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system 384-format

sequencing pipeline with 384 multiplexed capillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4 Kawai, T., Shibagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,

Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,

Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I.,

Saito, T., Okazaki, Y., Gojobori, T., Hara, H., Kasukawa, T., Saito, R.,

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, J.,

Fleischmann, W., Saastadland, J., Gissi, C., King, R., Kachiwagi, H.,

Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,

Quackenbush, J., Schrim, L. M., Staudt, F., Suzuki, K., Tomita, M.,

Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,

Baldarelli, R., Harsh, G., Hake, J., Hoffell, D., Hoshino, H.,

Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Holt, C.,

Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,

Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,

Marchionni, L., Mashima, J., Mazzarelli, J., McKaet, S., F., Nardone, F.,

Rao, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,

Sato, K., Schonbach, C., Scya, T., Shibata, Y., Storch, K. P., Suzuki, H.,

Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L.,

Wynshaw-Harris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,

and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection

Nature 404 (6821), 685-690 (2001)

21085660

11237851

5 (bases 1 to 1364)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,

Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanaoka, T.,

Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,

Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, I.,

Kato, M., Kawai, T., Kojima, Y., Konno, H., Kondo, M., Koyama, S.,

Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,

Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,

Saito, H., Saito, F., Sakai, C., Sakai, Y., Sano, H., Sasaki, D.,

Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,

Suzabe, Y., Suzuki, H., Tagami, M., Tazawa, A., Takahashi, F.,

Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamada, I.,

Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and

Hayashizaki, Y.

Hayashizaki, Y.

Direct Submission

Submitted (10 Jun 2000) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome



Genome version 5.1.4

Copyright (c) 1993 - 2003 CompuGen Ltd.

EM nucleic nucleic search, using SW model

Run on: January 16, 2003, 15:20:22 : Search time 2974.47 Seconds
(without alignments)
18469.506 Million cell updates/sec

Filter: US 09 026-459A 30

Perfect score: 4392

Sequence: 1 GGCACATGACATTTTACTGCT.....AAATGAGATTATGATGACT 3392

Scoring table: IDENTITY 90%

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 42406142

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estb01.*

2: em_estb01.*

3: em_estb01.*

4: em_estb01.*

5: em_estb01.*

6: em_estb01.*

7: em_estb01.*

8: em_estb01.*

9: qb_estb01.*

10: qb_estb01.*

11: qb_estb01.*

12: qb_estb01.*

13: qb_estb01.*

14: qb_estb01.*

15: em_estb01.*

16: em_estb01.*

17: qb_qss.*

18: em_qss.*

19: em_qss.*

20: em_qss.*

21: em_qss.*

22: em_qss.*

23: em_qss.*

24: em_qss.*

25: em_qss.*

26: em_qss.*

27: em_qss.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	910.2	26.8	1024	13	BM450031	BM450031 AGNCOURT
2	909.4	26.8	1051	13	BM561083	BM561083 AGNCOURT
3	879.8	25.9	1364	11	AK011246	AK011246 Mus muscu
4	856	26.2	1081	14	BQ439221	BQ439221 AGNCOURT
5	844.8	24.9	881	14	BQ420275	BQ420275 AGNCOURT
6	786	23.2	1001	13	BM475603	BM475603 AGNCOURT

7	752.4	22.2	1122	13	BM456881	BM456881 AGNCOURT
8	725.4	21.4	727	9	AL597811	AL597811 DKFZp4146
9	720.6	21.2	890	14	BQ222227	BQ222227 AGNCOURT
10	699.6	20.6	823	12	BG010661	BG010661 602611848
11	698.2	20.6	792	9	AL563757	AL563757 AGNCOURT
12	692.2	20.4	767	12	BQ570156	BQ570156 602611848
13	691.8	20.4	972	13	BM466194	BM466194 AGNCOURT
14	678	20.3	835	13	BM454724	BM454724 AGNCOURT
15	660	19.5	674	13	B1868459	B1868459 603492450
16	643.6	19.0	797	10	BE539278	BE539278 601064567
17	642.2	18.9	581	12	BG616219	BG616219 602611848
18	617	18.2	618	13	BM264009	BM264009 1434108.X
19	614	18.1	594	10	HE082846	HE082846 KC2 H1064
20	605.2	17.8	623	10	AW583181	AW583181 1411009.Y
21	605	17.8	605	9	AL599643	AL599643 DKFZp4146
22	596.4	17.6	922	12	BQ253543	BQ253543 602611848
23	580	17.1	580	9	AL599105	AL599105 DKFZp4146
24	575.2	17.0	914	9	AA763411	AA763411 VW54004.X
25	566.4	16.7	588	10	AW468234	AW468234 CM3 H1018
26	558	16.5	871	9	AA763485	AA763485 VW54101.X
27	539.2	15.6	598	10	AW582367	AW582367 140608.Y
28	530	15.6	549	9	AL642790	AL642790 DKFZp4146
29	528	15.6	572	9	AL120941	AL120941 DKFZp4146
30	516.8	15.2	597	12	HE931236	HE931236 MRO H1015
31	513.2	15.1	594	10	AV715333	AV715333 AV715333
32	512.2	15.1	734	12	BF144857	BF144857 601790443
33	505.2	14.9	736	13	B1151568	B1151568 602916108
34	502.4	14.8	832	9	AL646048	AL646048 VW54018.X
35	493.8	14.6	525	9	AA258255	AA258255 2159404.Y
36	486.4	14.3	584	12	HG149050	HG149050 udbrc01.Y
37	484.4	14.3	493	10	AW502887	AW502887 U1-HF HNO
38	481.4	14.2	484	10	HE168095	HE168095 QV3 H1051
39	480.2	14.2	659	13	B1821116	B1821116 603035591
40	476.6	14.1	2083	13	BM456728	BM456728 AGNCOURT
41	474.4	14.0	734	13	B1526982	B1526982 602923777
42	459.2	13.5	497	12	BP934784	BP934784 KC3 H1094
43	457	13.5	457	9	AL093215	AL093215 BQ04002.X
44	457	13.5	742	9	AL598706	AL598706 DKFZp4146
45	454.2	13.4	869	12	BE865749	BE865749 601678131

ALIGNMENTS

RESULT 1

BM450031

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AGNCOURT_6493544 NIH_MGC_72 Homo Sapiens cDNA clone IMAGE:5528037
5' mRNA sequence

BM450031 1024 bp mRNA linear EST 05-FEB-2002

BM450031 GI:18499071

Human

Homo Sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 1023)

NIH-MGC <http://mdc.ncbi.nlm.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: graphs@mail.nih.gov

Tissue Procurement: ATCC/DCTD/UTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Cloning Strategy: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: CLAM12304 row: 4 column: 22

High quality sequence stop: 643

Location/Qualifiers

1..1023

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528037"
/clone_lib="NIH_MGC_72"
/tissue_type="melanocytic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Skin; Vector: pCMV-Sport6; Site_1: Not I; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."
BASE COUNT      334 a      211 c      197 g      281 t
ORIGIN
Query Match.          26.8%; Score 910.2; DP 13; Length 1023;
Best local similarity 98.4%; Pred. No. 2 5e-165;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

QY 1040 CAACATCAAGCAATATACACCTTAACATCAAAATCAATGTAACAGCAATACAGCAATCTCC 2039
DB 1 CAACATCAAGCAATATACACCTTAACATCAAAATCAATGTAACAGCAATACAGCAATCTCC 60

QY 2040 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 2099
DB 61 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 120

QY 2100 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 2159
DB 121 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 180

QY 2200 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 2219
DB 181 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 240

QY 2220 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 2279
DB 241 TATGTTGTTTCAGAGAGATTCAAAGTGTGTTTGATCAAGAGAGAGAGATGATGATCTAT 300

QY 2280 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2339
DB 301 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 360

QY 2340 AGTATCAATATGTTGATATATATATATATATATATATATATATATATATATATATATAT 2399
DB 461 AGTATCAATATGTTGATATATATATATATATATATATATATATATATATATATATATAT 420

QY 2400 ATGTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2459
DB 481 ATGTTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 540

QY 2500 GAAAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2519
DB 541 GAAAAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 600

QY 2580 AAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2639
DB 601 AAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 660

QY 2640 GATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2699
DB 661 GATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 720

QY 2700 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2759
DB 721 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 780

QY 2760 AAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2817
DB 781 AAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 840

```

```

QY 2818 GTTATTTATATACAGATTCGAAATCTTGTGTAATATCTTGTGTAATATCTTGTGTAATATCT 2877
DB 841 GTTATTTATATACAGATTCGAAATCTTGTGTAATATCTTGTGTAATATCTTGTGTAATATCT 900

QY 2878 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2933
DB 901 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 960

QY 2934 AGATTTGATTTTGTGTAATATCTTGTGTAATATCTTGTGTAATATCTTGTGTAATATCTTGT 2993
DB 961 AGATTTGATTTTGTGTAATATCTTGTGTAATATCTTGTGTAATATCTTGTGTAATATCTTGT 1020

QY 2994 A 2994
DB 1021 A 1021

RESULT 2
BM561083
LOCUS
DEFINITION
  BM561083 1051 bp mRNA linear 181 20-SEP-2002
  AGENCOURT_6566133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550905
  5' mRNA sequence.
ACCESSION
  BM561083
VERSION
  BM561083.1 GI:18896043
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Gracilata, Vertebrata, Euteleostomi;
  Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
  1 (bases 1 to 1051)
  NIH-MGC http://mgi.nih.gov/.
AUTHORS
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
JOURNAL
  Contact: Robert Strausberg, Ph.D.
  Email: eqabbs@mail.nih.gov
COMMENT
  Tissue procurement: ATCC
  cDNA library preparation: Life Technologies, Inc.
  cDNA library arrayed by: the I.M.A.G.E. Consortium (LLNL)
  DNA sequencing by: Agencourt Bioscience Corporation
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LAM12263 row: m column: 18
  High quality sequence stop: 737.
FEATURES
  source
    location/Qualifiers
    1..1051
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:5550905"
    /clone_lib="NIH_MGC_67"
    /tissue_type="retinoblastoma"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: eye; Vector: pCMV-Sport6; Site_1: Not I; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."
  BASE COUNT      330 a      221 c      197 g      299 t
  ORIGIN
    Query Match          26.8%; Score 909.4; DP 13; Length 1051;
    Best local similarity 98.4%; Pred. No. 3 5e-165;
    Matches 948; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1138 AAAATGTTGAGCAAGGTTTGTGTTGAGGAAATGGATATCAAGAGATTAAGATTCGAGTTTCG 1197
DB 1 ATATGTTGAGCAAGGTTTGTGTTGAGGAAATGGATATCAAGAGATTAAGATTCGAGTTTCG 60

QY 1198 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1257
DB 61 TTTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 120

QY 1258 AATTTAGCAAACTTCGAGATGACACATTTTCATATCTTTATATCTGCTGCTTTT 1317
DB 1317 AATTTAGCAAACTTCGAGATGACACATTTTCATATCTTTATATCTGCTGCTTTT 1377

```

JOURNAL MEDLINE PUBMED REFERENCE	TITLE
99279253 10349636	Meth. Enzymol. 403, 19-44 (1999)
2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.: Normalization and subtraction of cap trapped selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1640 (2000)
20499474 11042159	RIKEN integrated sequence analysis (RISA) system: 484-format sequencing pipeline with 484 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
3	Shibata, K., Itoh, M., Aizawa, K., Nakajima, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsumaru, T., Yoshino, M., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikeda, T., Kashiwagi, K., Fujikake, S., Inoue, K., Todaka, Y., Izawa, M., Ohara, E., Watanabe, K., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawaji, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.: RIKEN integrated sequence analysis (RISA) system: 484-format sequencing pipeline with 484 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
20530913 11076861	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, K., Saito, T., Okazaki, Y., Gotohori, T., Ikeno, H., Kasai, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, J., Fleisclmann, W., Gaasterland, F., Glass, C., King, R., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, J. M., Staehli, F., Suzuki, K., Tomita, M., Wanier, L., Washio, T., Sakai, K., Oikido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Bollen, D., Brownstein, M., Bull, P., Carninci, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Humé, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Montanari, F., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schaubach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wysslow, R., Yoda, A., Yoshida, K., Hasegawa, Y., Kawajiri, K., Koh, S., and Hayashizaki, Y.: Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660 11217851	5 (bases 1 to 164)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arab, A., Arakawa, T., Baldarelli, R., Bonito, H., Brownstein, M., Bull, P., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Harada, A., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirao, T., Hori, F., Humé, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawaji, J., Kojima, Y., Kono, H., Konda, M., Koyama, S., Kurihara, C., Matsuyama, I., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ogiwara, M., Okazaki, Y., Oikido, T., Ohta, T., Quackenbush, J., Saito, H., Saito, P., Sakai, Y., Sakai, Y., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tadawa, A., Takahashi, F., Tanaka, I., Tejima, Y., Toya, I., Yamamura, I., Yamataka, J., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.: Direct Submission	
5,6,10,14 (20 Feb 2000)	Yoshida, K., Hayashizaki, Y.: The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suchoiro cho, Tsurumi ku, Yokohama, Kanagawa 230 0045, Japan (E mail: yoshi@genome-res.riken.ac.jp, URL: http://genome.qsc.riken.go.jp/, Tel: 81 45 503 9222, Fax: 81 45 503 9216)
please visit our web site (http://www.riken.ac.jp/) for further details.	
JOURNAL MEDLINE PUBMED REFERENCE	TITLE
20530913 11076861	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, K., Saito, T., Okazaki, Y., Gotohori, T., Ikeno, H., Kasai, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, J., Fleisclmann, W., Gaasterland, F., Glass, C., King, R., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, J. M., Staehli, F., Suzuki, K., Tomita, M., Wanier, L., Washio, T., Sakai, K., Oikido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Bollen, D., Brownstein, M., Bull, P., Carninci, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Humé, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Montanari, F., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schaubach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wysslow, R., Yoda, A., Yoshida, K., Hasegawa, Y., Kawajiri, K., Koh, S., and Hayashizaki, Y.: Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660 11217851	5 (bases 1 to 164)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arab, A., Arakawa, T., Baldarelli, R., Bonito, H., Brownstein, M., Bull, P., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Harada, A., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirao, T., Hori, F., Humé, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawaji, J., Kojima, Y., Kono, H., Konda, M., Koyama, S., Kurihara, C., Matsuyama, I., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ogiwara, M., Okazaki, Y., Oikido, T., Ohta, T., Quackenbush, J., Saito, H., Saito, P., Sakai, Y., Sakai, Y., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tadawa, A., Takahashi, F., Tanaka, I., Tejima, Y., Toya, I., Yamamura, I., Yamataka, J., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.: Direct Submission	
5,6,10,14 (20 Feb 2000)	Yoshida, K., Hayashizaki, Y.: The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suchoiro cho, Tsurumi ku, Yokohama, Kanagawa 230 0045, Japan (E mail: yoshi@genome-res.riken.ac.jp, URL: http://genome.qsc.riken.go.jp/, Tel: 81 45 503 9222, Fax: 81 45 503 9216)
please visit our web site (http://www.riken.ac.jp/) for further details.	
JOURNAL MEDLINE PUBMED REFERENCE	TITLE
20530913 11076861	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, K., Saito, T., Okazaki, Y., Gotohori, T., Ikeno, H., Kasai, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, J., Fleisclmann, W., Gaasterland, F., Glass, C., King, R., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, J. M., Staehli, F., Suzuki, K., Tomita, M., Wanier, L., Washio, T., Sakai, K., Oikido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Bollen, D., Brownstein, M., Bull, P., Carninci, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Humé, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Montanari, F., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schaubach, C., Seva, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wysslow, R., Yoda, A., Yoshida, K., Hasegawa, Y., Kawajiri, K., Koh, S., and Hayashizaki, Y.: Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
21085660 11217851	5 (bases 1 to 164)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arab, A., Arakawa, T., Baldarelli, R., Bonito, H., Brownstein, M., Bull, P., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Harada, A., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirao, T., Hori, F., Humé, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawaji, J., Kojima, Y., Kono, H., Konda, M., Koyama, S., Kurihara, C., Matsuyama, I., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ogiwara, M., Okazaki, Y., Oikido, T., Ohta, T., Quackenbush, J., Saito, H., Saito, P., Sakai, Y., Sakai, Y., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tadawa, A., Takahashi, F., Tanaka, I., Tej	

QY 1091 GTTAAAGAAATCTGATT CTATTTTAAAGATGGACAGTGAATGCAAAAGAAAGATAT 1088
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 824 CTGAAAAAATGGAATTTTATTTTAAACAGTGGACGGCAATTCACAAAGAAAGC 882
 QY 1089 ATGAAAAAGATGAGGATATAGAGATACATCTTAA 1124
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 884 TATATGGAAGAGAGGATAGAGATATATAGATAA 918
 REFSeq 7
 LOCUS BM545881 1122 bp mRNA Linear EST 20 FEB 2002
 DEFINITION AGENE:GRT.5595184 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5588245
 5', mRNA sequence.
 ACCESSION BM545881
 VERSION 1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1122)
 AUTHORS NIH-MGC <http://mrc.nhlbi.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cdp@bbs.tamh.nih.gov
 Issue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Sequencing by: Applied Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the T.M.A.G.E. consortium/LINL at:
<http://tmage.liv.edu>
 Plate: L1AM12459 Row: 4 Column: 14
 High quality sequence start: 48
 High quality sequence stop: 773.
 FEATURES
 Location/Qualifiers
 1..1122
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5588245"
 /zelle="NIH-MGC_125"
 /lab_host="DB10B"
 /note="Ovarian ovary (pool of 3); Vector: pCMV-Sport6;
 Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool
 of three ovaries, from females ranging in age from 38 to
 49 yo. Library is oligo-dT primed and directionally cloned
 (EcoRV site is destroyed upon cloning). Average insert,
 size 2.1 kb, insert size range 1-3.5 kb. Library is
 normalized and enriched for full-length clones and was
 constructed by C. Gruber (Invitrogen). Research Genetics
 tracking code 036."

DB 420 CAGAAATGGAGAAATGATTTACATGAAACAGAAAGAAAAAGAAATGAT 479
 QY 2599 AGCATGATACCTCAACACAGCAAGCAAGACAGATCTAGATCTTCTGATCAATG 2658
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 480 AGCATGATACCTCAACACAGCAAGCAAGACAGATCTAGATCTTCTGATCAATG 549
 QY 2659 GTATAGCTCTGATTCATTCCTCTCAACATGATGATGATGATGATGAT 2718
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 540 GTACACCTCTGATTCATTCCTCTCAACATGATGATGATGATGATGAT 599
 QY 2719 TTATGGGCAATTTTATTAATCTCTAGCTCTTTTGTGGATATATAATGATGAT 2778
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 600 TTATGGGCAATTTTATTAATCTCTAGCTCTTTTGTGGATATATAATGATGAT 659
 QY 2779 GTTGGGTCATTCCTCAAGTCACATGCAAAAGTACATGATGATGATGATGAT 2838
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 560 GTTGGGTCATTCCTCAAGTCACATGCAAAAGTACATGATGATGATGATGAT 719
 QY 2839 ATCTGTGTAAGTCCTGATTTTAAAAAAGTGTGATGATGATGATGATGATGAT 2898
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 720 ATCTGTGTAAGTCCTGATTTTAAAAAAGTGTGATGATGATGATGATGATGAT 779
 QY 2899 AATCTGTGCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2958
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 780 AATCTGTGCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 839
 QY 2959 CTCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3018
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 840 CTCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 898
 QY 3019 TAATTATA--TGATATATTTTAAATTAATGATGATGATGATGATGATGAT 3075
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 860 TAATTATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
 QY 3076 TGTATCTTCCAAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3135
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 959 TGTATCTTCCAAATGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1017
 QY 3136 TGTCAAAAT--GATATATTTACAAATAGAAAAATTTAAATTTTAAATTTTAAATTT 3194
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 1018 TGTCAAAATGCAATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
 QY 3195 TATTTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3259
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 1078 TATTTTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
 RESULT 8
 AL597811
 LOCUS
 DEFINITION DKEZp41300816_r1 413 (synonym: hlee2) Homo sapiens cDNA clone
 DKEZp41300816 5', mRNA sequence.
 ACCESSION AL597811
 VERSION 1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 727)
 AUTHORS Poustka A., Wellenreuther R., Mewes H.W., Well H. and Wiemann S.).
 TITLE EST (Poustka A., Wellenreuther R., Mewes H.W., Well H. and Wiemann S.).
 JOURNAL Unpublished (1999)
 COMMENT Contact: Poustka A.J.
 Department Lebach
 Max Planck Institute for Molecular Genetics
 Ihnestrasse 74, 14195 Berlin, Germany
 Tel: +49-30-84131623
 Fax: +49-30-84131128
 Email: poustka@mpi-lebach.dahlem.mpg.de
 This is the 5' sequence of the clone insert

Query Match
 Best Local Similarity 72.2% Score 752.4; DB 13; Length 1122;
 Matches 827; Conservative 0; Mismatches 42; Indels 6; Gaps 5;
 QY 2459 TTGGGATTTTGAAGATTCAGAAAAATTAATGATGATGATGATGATGATGATGATG 2418
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 240 TTGGGATTTTGAAGATTCAGAAAAATTAATGATGATGATGATGATGATGATGATG 299
 QY 2419 CTCAAAATTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2478
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 400 CTCAAAATTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 359
 QY 2479 ATTGAAGATTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2538
 DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
 DB 460 ATTGAAGATTAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 419
 QY 2539 CAGAAATGCGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2598

Query Match	20.4%	Score	692.0	EB	12	Length	767
Best Local Similarity	9.1%	Freq.	86	Size	137		

Query Match 19.5%; Score 660; Db 13; Length 672;
Best Local Similarity 99.9%; Pred. No. 4,16-117;
Matches 671; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2258 ATATTTCACTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2317
DB 1 ATATTTCACTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 60

QY 2318 TGACTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2377
DB 1 TGACTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 120

QY 2378 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2437
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 180

QY 2438 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2497
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 240

QY 2498 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2557
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 300

QY 2558 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2617
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 360

QY 2618 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2677
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 420

QY 2678 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2737
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 480

QY 2738 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2797
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 540

QY 2798 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2856
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 600

QY 2857 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2916
DB 1 TGCTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 660

QY 2917 ATACTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 2976
DB 1 ATACTGTAAGAGTCATATAAAATTTGAGAGGTCGTGAAACACCAACAAAA 672

Search completed: January 18, 2003, 04:06:59
Job time: 4001.45 secs




```

Db 544 CCAAGAGAAAAAGCTTAAATAGATGCTAAATTTATATGAAAGAGAGACAAAGCA 603
QY ThrSerAlaPheGlnThrGlnIysProLeuLysSerThrSerLeuSerLeuPheTyrLys 548
Db 545 AACTCAAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 553
QY LysValTyrArgLeuAlaTyrLeuArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 618
Db 546 AAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
QY HisProGlnLeuGlnIleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 638
Db 724 CACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 783
QY GlnLeuGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 652
Db 784 GAACTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 843
QY LysValLysAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 678
Db 844 AAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 903
QY GlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 697
Db 904 CACGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 963
QY IleLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 714
Db 964 TATAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020

RESULT 4
LOCUS BQ220275 881 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENTOURT 7572621 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6044522
5' mRNA sequence
ACCESSION BQ220275
VERSION BQ220275.1 GI:20401675
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
EXTRACTOR: MRC Gen. Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 881)
NIH-MGC http://mgs.nhl.nih.gov/.
NATIONAL INSTITUTES OF HEALTH, Mammalian Gene Collection (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
EMAIL: rstrauss@nhi.nih.gov
Tissue: Prostate, ATCC
cDNA library: Preparation, Life Technologies, Inc.
cDNA library: Arrayed by the I.M.A.G.E. Consortium (LBN)
cDNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: LLAM3287 row: C column: 03
High quality sequence stop: 690
Location/Qualifiers
1. 881
/organism="Homo sapiens"
/clone="IMAGE:6044522"
/clone="NIH_MGC_92"
/tissue_type="embryonal; carcinoma; cell line"
/lab_host="293T6 (plaque resistant)"
/notes="organ. testis; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 413 a 149 c 155 g 264 t
ORIGIN
BM475603

```

```

Alignment Scores: 2 076-145 Length: 881
Pred No 1448.00 Matches: 288
Percent Similarity: 98.30% Conservative: 1
Best Local Similarity: 97.96% Mismatches: 3
Query Match: 32.15% Indels: 2
DB: 14 Gaps: 0
US-09-026-459a-31 (1 874) X BQ220275 (1-881)

```

```

QY 154 MetGluAspAspLeuValIleSerPheGlnLeuMetLeuLysValLeuAspTyrPheIle 173
Db 3 ATGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 62
QY 174 LysLeuSerGlnGlnGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 193
Db 63 AAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 122
QY 194 GlySerProArgThrProArgArgGlyGluAsnArgSerAlaArgIleAlaLysGlnLeu 213
Db 123 GCTTCACCTGCAATACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 192
QY 214 GluAspAspThrArgIleLeuLeuValLeuLeuLysLysLysLysLysLysLysLys 233
Db 193 GAAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 242
QY 234 ValLysAsnValTyrLeuLysAsnPheLeuLeuPheMetAsnSerLeuLeuValThr 253
Db 243 GTTAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 402
QY 254 SerAspGlyLeuProGlnValGluAsnLeuSerLysArgTyrGlnGlnLeuLeuLys 273
Db 303 CTAAAGCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 362
QY 274 AsnLysAspLeuAspAlaArgLeuPheLeuAspHisAspLysThrLeuGlnThrAspSer 293
Db 363 ATATAGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 422
QY 294 IleAspSerPheGlnThrGlnArgThrProArgLysSerAsnLeuAspGlnLeuValAsn 313
Db 423 ATAGCAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 482
QY 314 ValIleProPheThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMet 333
Db 483 GTAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 542
QY 334 IleLeuAsnSerAlaSerAspGlnProSerGluAsnLeuLeuLeuLeuLeuLeuLeu 353
Db 543 ATTTTAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 602
QY 354 ThrValAsnProLysLeuSerIleLeuLysArgValLysAspIleGlyTyrIlePheLys 373
Db 603 AATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 662
QY 374 GlnLysThrAlaLysArgValGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 393
Db 663 CAGAAATTTCTTAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 722
QY 394 LeuGlyValArgLeuTyrTyrArgValMetGluSerMetLeuLysSerGlnGlnLeuArg 413
Db 723 CTGAGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 782
QY 414 LeuSerIleLeuAsnPheSerLysLeuLeuAsnAspAsnIlePheHisMetSerLeu 433
Db 783 TATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 842
QY 433 GAlaGlnAlaLeuGlnValValMetAlaThrTyrSer 445
Db 843 TGGTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 880

```

RESULT 4
BM475603

LOCUS BM475603 1001 bp mRNA Linear EST 05-FEB-2002
 DEFINITION ACEN004041 6480007 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575683
 5' mRNA sequence.
 ACCESSION BM475603
 VERSION BM475603.1 GI:18524645
 KEYWORDS EST.
 ORGANISM human.
 REFERENCE 1 (bases 1 to 1001)
 AUTHORS NIH MGC http://mgi.nch.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rstraus@nch.nih.gov
 Tissue Procurement: ATCC
 cDNA library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequenced by: Apocorty Bioscience Corporation
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plates: BM475603, row: 1, column: 04
 High quality sequence start: 17
 High quality sequence stop: 734.
 Location/Qualifiers
 1..1001
 Zonedatum "Homo sapiens"
 Zonedatum "axons:9606"
 Zonedatum "BM475603"
 Zonedatum "NIH MGC 92"
 Zonedatum "Embryonal carcinoma, cell line"
 Zonedatum "U937 (phage resistant)"
 Zonedatum "Vector: pCMV-Sport6; Site 1: Not I;
 Site 2: Sal I; cloned unidirectionally; cDNA-3' primer
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by life technologies.
 Note: This is a NIH-MGC Library."
 BASE COUNT 449 a 176 c 186 g 288 t
 ORIGIN
 Alignment Scores:
 Prod. No.: 1480-143 Length: 1001
 Score: 1499.00 Matches: 281
 Percent Similarity: 94.87% Conservative: 10
 Best Local Similarity: 90.65% Mismatches: 10
 Query Match: 29.74% Indels: 9
 Gaps: 3
 03-09-026-459a-31 (1-874) x BM475603 (1-1001)
 97 81 LeuLysGluValAspThrSerThrLysValAspAlaMetSerArgLeuLeuLysLys 100
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 10 81 GCGCGCAAGAAATGCAATCAAGAAATGCAATCAAGAAATGCAATCAAGAAATGCAAG 97
 97 101 TyrAspValLeuPheAlaLeuPheSerLysLeuGluArgThrGlyLeuLeuLeuLeu 120
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 10 98 TATGATGTAATGTTTGTGATGTTTGTGATGTTTGTGATGTTTGTGATGTTTGTGATGTT 157
 97 120 GthGlnProSerSerLeuSerThrGlnLeuAspSerAlaValLeuLysVal 140
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 10 158 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 217
 97 140 TTrpThrPheLeuAlaLysGlyGlyValLeuGlnMetGlnAspLeuVal 160
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 10 218 TGGATCAAT 277
 97 160 eSerPheLeuMetLeuLysValLeuAspThrPheLeuLysLeuSerProMetLe 180
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 10 278 TTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 337
 97 180 AlaLysGluProGlyThrAlaValLeuProThrAsnGlySerProArgThrProArg 200

DB 338 GGTCAAGAACCAATATAAAACAGCTGTATATACCATTAATGCTTCAATCAAAATCCAG 497
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 200 GArqLysGlnSerArgSerAlaArgAlaLeuLysGlnGlnGlnAspThrArgAla 220
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 398 GCGAGTCAGAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 457
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 220 cSLaValLeuGlySerLysThrGlnSerLysThrGlnSerLysThrGlnSerLys 240
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 458 TGAAGTTCCTGCTGAAGAACATGATGATGATGATGATGATGATGATGATGATGAT 517
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 240 sAsnPheLeuProPheMetAsnSerLeuSerLysValLeuSerLysValLeuSer 260
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 518 AAATTTTATAGCTTTTATGAATTCCTGACATGATGATGATGATGATGATGATGAT 577
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 260 LGLuAsnLeuSerLysArgTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 280
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 578 TGAAGAACCTCTAAACATACAGAGAAATTTATCTTAAATTAATTAATTAATTAAT 637
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 280 GlnPheLeuAspHisAspLysThrLeuGlnThrAspSerLeuAspSerLeuThrGln 300
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 638 ATTATTTTGGATCATGATATAAACTCTTCAGACTTATTTATATATATATATATAT 697
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 300 rArqThrProArgLysSerAsnLeuAspGlnValAsnValLeuProThrHisThrPr 320
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 698 GACAACACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 757
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 320 cValArgThrValMetAsnThrLeuGlnGlnLeuMetMetLeuAsnSerAlaSerAs 340
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 758 AGTAGGACGTGATAC 817
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 340 pGlnProSerGlnAsnLeuLeuSerTyrPheAsnAspGlySerThrValAsnThrGly 359
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 818 TCAAGCTTCGAAATGATGATTTTCTATTTTAAACAAATGACGCTGGAATTCGAAAG 877
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 359 uSerLeuLysArgGlyValLysAspThrGlyValLysAspThrGlyValLys 375
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 878 AAAGCTATACCGGAAAGAAACCGGTAAGCAATAGCAATAACCTCTCTTTTAAAGAAA 937
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 QY 376 --PheAlaLysAlaValGly 381
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 DB 938 TTCCTGCTAAAAGCTGCGG 957
 100 CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC CCCCC
 RES011 5
 AL597811 727 bp mRNA Linear EST 14 AUG 2001
 LOCUS DKEZp1406816_r1 313 (5702236, Fluc2) Homo sapiens cDNA clone
 DEFINITION DKEZp1406816 5' mRNA sequence.
 ACCESSION AL597811
 VERSION AL597811.1 GI:15160502
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 727)
 AUTHORS Poustka A., Wellenreuther R., Mewes H.W., Weil B. and Wiemann S.
 TITLE EST (Poustka A., Wellenreuther R., Mewes H.W., Weil B. and Wiemann S.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Poustka A.J.
 Department Lebach
 Max Planck Institute for Molecular Genetics
 Thiersstrasse 74, 14195 Berlin, Germany
 Tel.: +49-30-84141624
 Fax: +49-30-84141128
 Email: poustka@mpg-berlin.dahlem.mpg.de
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
 Heidelberg/Germany within the cDNA sequencing consortium of the
 German Genome Project.

with a Not I oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pTZ19 vector. Library is normalized. Library was constructed by Benito Soares and M. Fatima Bonaldo.

BASE COUNT	257 a	178 c	217 g	261 t	1 g-fs
ORIGIN					
Alignment Scores:					
Prod. No.:	4.1e-103	Length:	914		
Score:	1056.50	Matches:	234		
Percent Similarity:	82.14%	Conservative:	19		
Best local Similarity:	75.97%	Mismatches:	53		
Query Match:	23.46%	Indels:	6		
DB:	9	Gaps:	2		
US-09-026-459a-31 (1-874) x AA763411 (1-914)					
QY	282	PhcGlnAspHisAspLysThrLeuGlnThrAspSerThrAspSerPheGlnThrGlnArg	301		
DB	914	TTTTGGATGCAATAAAGCGTCGAGGTGT-TGCTTTAGGCTTTTTTAAAGCGGAG	857		
QY	302	ThrProAcTysSerAsnLeuAspGluValAsnValIleProProHisThrProVal	321		
DB	856	AGGACGCAAAACACCTTGT-CAAGACGCAAAACGCGGTACTGCGCACCCCGCTT	798		
QY	322	ArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGln	341		
DB	797	TGATGTGTTTGAAAGTATGCAACAAATTTATGATTTTAAACGCGGATGAGCGG	736		
QY	342	ProSerGlnAsnLeuLeuSerTyrPheAsnAsnGlyThrValAsnProLysGlnSerIle	361		
DB	747	CGCTCAGAAINTGCTT-TGCTGCTTAATATGAGCGCGGATAGCGCAAGAAATGCG	679		
QY	362	IlePysArgValLysAspIleLeuTyrIlePheGlnLysPheAlaLysAlaValGly	381		
DB	678	CTAAAGCGGCTAAGACGCTGGGACACACACCTTCTCTA-ACCGCTGCG	623		
QY	382	GlnGlyGysValGlnIleLeuGlySerGlnArgTyrLysLeuGlyValAlaLeuTyrTyrArg	401		
DB	622	CGAGGCTGTCGACATCGGATACGAGGATAGGATATAAATTTTGAGTGGCATGTATACGGT	563		
QY	402	ValMetGlnSerMetLeuLysSerGlnGlnArgGlnSerIleGlnAsnProSerLys	421		
DB	562	GTCATGGAATCATGCTATGATACAGAAAGAAAGTTTGTTCATTCAGAAATTTAGCAA	503		
QY	422	LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaLysAlaLeuGlnValValMet	441		
DB	502	CTGTAAAGAGAAATGCAATCTTACAGGCTGTGCTCTGAGAGTGTAAAG	443		
QY	442	AlaIleTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro	461		
DB	442	GTATGATATAGCAAGATCATTTGACATGATTCATTCGAAAGATTTGTCTTCTGGG	383		
QY	462	TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGlnSerPhe	481		
DB	382	TGATCTCAAGCTACTTAATTTAAAGAGCTTTGATTTTTCAGAAAGTGTATGAAATTTT	323		
QY	482	IleLysAlaGlnLysIleAsnLeuThrArgGlnMetIleLysIleLeuGlnArgCysGlnHis	501		
DB	322	ATCAAAAGCAAGCAAAATCTCAAGCAAGAAAGATAGAAACAACTAGAAAGCAATGACAT	263		
QY	502	ArgIleMetGlnSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGln	521		
DB	262	CAATATAGGAAGCTGTATAGGCTTTTCAGATTCACCTTTATTGTGATCTATTAGCAG	203		
QY	522	SerLysAspArgGlnGlyProThrAspHisLeuGlnSerAlaCysProLeuAsnLeuPro	541		
DB	202	TCAAAGATGAGCAAGCAAGT-CATAAGCTGAAAGCTGCTGCTCAGAGCTGCT	146		
QY	542	LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys	561		
DB	146	CTCAAGCAAGCAAGCAAGTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	86		

QY	562	LysCysLysThrThrArgValAsnSerThrAlaAsnAlaLeuThrGlnAlaThrSerAla	581		
DB	85	AGAAATTCAGATCAAGAGTGTAAATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	26		
QY	582	PheGlnThrGlnLysProLeuLys	589		
DB	25	TTCCATACACAGAGCAATGAAA	2		
RESULT 10					
LOCUS	AL564757	792 bp	mRNA	Linear	EST 16 FEB 2001
DEFINITION	AL564757 L11_NF1001_N804 Homo sapiens cDNA clone (S000073A05)				
ACCESSION	AL564757				
VERSION	AL564757.1	GI:12914464			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
TITLE	1 (bases 1 to 792)				
JOURNAL	Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.				
COMMENT	Full-length cDNA libraries and normalization Unpublished (2001) Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 Evry cedex - France Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr. Location/Qualifiers				
FEATURES	1..792				
SOURCE	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="GSD00073A05" /clone_lib="L11_NF1001_N804" /sex="male" /tissue_type="neuroblastoma cells" /lab_host="DH10B" /note="Argan, brain. Vector: pMV360ki 6, 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pMV360ki 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Genoscope Technologies, a division of Invitrogen 4860 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: lliang@litech.com URL: http://fulllength.invitrogen.com"				
BASE COUNT	246 a	148 c	117 g	255 t	26 others
ORIGIN					
Alignment Scores:					
Prod. No.:	9.4e-102	Length:	792		
Score:	1043.00	Matches:	217		
Percent Similarity:	84.54%	Conservative:	7		
Best local Similarity:	81.89%	Mismatches:	38		
Query Match:	23.16%	Indels:	4		
DB:	9	Gaps:	0		
US-09-026-459a-31 (1-874) x AL564757 (1-792)					
QY	186	LysThrAlaValIleThrThrLeuAsnGlySerProAlaThrProAlaGlnGlyGlnAsnArg	205		
DB	791	AAAGAGCTTATACCATTAATGTTCACTCGAAACCGACGAGGTACAAAGG	742		
QY	206	SerAlaArgGlnAlaLysGlnLeuGlnAsnAspThrArgIleIleThrValLeuGlyLys	225		
DB	731	ACTGACCGGTACCAACACAAATACAAATTCAGATTCCTGAAAA	672		
QY	226	GlnHisGlnCysAsnIleAspGlnValLysAsnValTyrPheLysAsnThrProPhe	245		
DB	671	GAACATGAATGTATATAGTCAGCTGAGAAATGCTTTCACAAAAATTAATAATCTTT	612		


```

KEYWORDS EST
SOURCE Homo sapiens
ORGANISM
REFERENCE
AUTHORS
    Iwasaki, N., Garcia-Correa, P., Vorjovski-Almeida, S., Briones, M. R.,
    Nishi, M., A. da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F.,
    Goldman, G. R., Carvalho, A. F., Matsukuma, A., Kaya, G. S., Simpson, D. H.,
    Brunstein, A., de Oliveira, P. S., Bucher, P., Jordeiro, C. V., O'Hare
    M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J., and
    Simpson, A. J.
TITLE
    Shotgun sequencing of the human transcriptome with OPP expressed
    sequence tags
JOURNAL
    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE
    20202653
COMMENT
    Contact: Simpson A.J.G.
    Laboratory of Cancer Genetics
    Ludwig Institute for Cancer Research
    Rua Prof. Antonio Trudente 109, 4 andar, 01509-010, Sao Paulo-SP,
    Brazil
    Tel: +55 11 2704922
    Fax: +55 11 2707001
    Email: asimpson@ludwig.org.br
    This sequence was derived from the FAPESP/LICK Human Cancer Genome
    Project. This entry can be seen in the following URL
    (http://www.ludwig.org.br/scripts/getform2.pl?1-459a-31)
    200 014 c06a43-2000-02-21a44-1)
    Seq primer: puc 18 forward
    High quality sequence stop: 58
    High quality sequence stop: 669.
FEATURES
    Location/Qualifiers
        1..694
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone_lib="BT0642"
            /dev_stage="Adult"
            /note="A mini-library was made by cloning products derived
            from ONESIES PCR (U.S. Letters Patent application No. 196
            716) Ludwig Institute for Cancer Research) profiles
            into the pUC 18 vector. Reverse transcription of tissue
            mRNA and cDNA amplification were performed under low
            stringency conditions."
        BASE COUNT 211 4 129 3 127 9 226 1 1
        ORIGIN
            1
            2
            3
            4
            5
            6
            7
            8
            9
            10
            11
            12
            13
            14
            15
            16
            17
            18
            19
            20
            21
            22
            23
            24
            25
            26
            27
            28
            29
            30
            31
            32
            33
            34
            35
            36
            37
            38
            39
            40
            41
            42
            43
            44
            45
            46
            47
            48
            49
            50
            51
            52
            53
            54
            55
            56
            57
            58
            59
            60
            61
            62
            63
            64
            65
            66
            67
            68
            69
            70
            71
            72
            73
            74
            75
            76
            77
            78
            79
            80
            81
            82
            83
            84
            85
            86
            87
            88
            89
            90
            91
            92
            93
            94
            95
            96
            97
            98
            99
            100
            101
            102
            103
            104
            105
            106
            107
            108
            109
            110
            111
            112
            113
            114
            115
            116
            117
            118
            119
            120
            121
            122
            123
            124
            125
            126
            127
            128
            129
            130
            131
            132
            133
            134
            135
            136
            137
            138
            139
            140
            141
            142
            143
            144
            145
            146
            147
            148
            149
            150
            151
            152
            153
            154
            155
            156
            157
            158
            159
            160
            161
            162
            163
            164
            165
            166
            167
            168
            169
            170
            171
            172
            173
            174
            175
            176
            177
            178
            179
            180
            181
            182
            183
            184
            185
            186
            187
            188
            189
            190
            191
            192
            193
            194
            195
            196
            197
            198
            199
            200
            201
            202
            203
            204
            205
            206
            207
            208
            209
            210
            211
            212
            213
            214
            215
            216
            217
            218
            219
            220
            221
            222
            223
            224
            225
            226
            227
            228
            229
            230
            231
            232
            233
            234
            235
            236
            237
            238
            239
            240
            241
            242
            243
            244
            245
            246
            247
            248
            249
            250
            251
            252
            253
            254
            255
            256
            257
            258
            259
            260
            261
            262
            263
            264
            265
            266
            267
            268
            269
            270
            271
            272
            273
            274
            275
            276
            277
            278
            279
            280
            281
            282
            283
            284
            285
            286
            287
            288
            289
            290
            291
            292
            293
            294
            295
            296
            297
            298
            299
            300
            301
            302
            303
            304
            305
            306
            307
            308
            309
            310
            311
            312
            313
            314
            315
            316
            317
            318
            319
            320
            321
            322
            323
            324
            325
            326
            327
            328
            329
            330
            331
            332
            333
            334
            335
            336
            337
            338
            339
            340
            341
            342
            343
            344
            345
            346
            347
            348
            349
            350
            351
            352
            353
            354
            355
            356
            357
            358
            359
            360
            361
            362
            363
            364
            365
            366
            367
            368
            369
            370
            371
            372
            373
            374
            375
            376
            377
            378
            379
            380
            381
            382
            383
            384
            385
            386
            387
            388
            389
            390
            391
            392
            393
            394
            395
            396
            397
            398
            399
            400
            401
            402
            403
            404
            405
            406
            407
            408
            409
            410
            411
            412
            413
            414
            415
            416
            417
            418
            419
            420
            421
            422
            423
            424
            425
            426
            427
            428
            429
            430
            431
            432
            433
            434
            435
            436
            437
            438
            439
            440
            441
            442
            443
            444
            445
            446
            447
            448
            449
            450
            451
            452
            453
            454
            455
            456
            457
            458
            459
            460
            461
            462
            463
            464
            465
            466
            467
            468
            469
            470
            471
            472
            473
            474
            475
            476
            477
            478
            479
            480
            481
            482
            483
            484
            485
            486
            487
            488
            489
            490
            491
            492
            493
            494
            495
            496
            497
            498
            499
            500
            501
            502
            503
            504
            505
            506
            507
            508
            509
            510
            511
            512
            513
            514
            515
            516
            517
            518
            519
            520
            521
            522
            523
            524
            525
            526
            527
            528
            529
            530
            531
            532
            533
            534
            535
            536
            537
            538
            539
            540
            541
            542
            543
            544
            545
            546
            547
            548
            549
            550
            551
            552
            553
            554
            555
            556
            557
            558
            559
            560
            561
            562
            563
            564
            565
            566
            567
            568
            569
            570
            571
            572
            573
            574
            575
            576
            577
            578
            579
            580
            581
            582
            583
            584
            585
            586
            587
            588
            589
            590
            591
            592
            593
            594
            595
            596
            597
            598
            599
            600
            601
            602
            603
            604
            605
            606
            607
            608
            609
            610
            611
            612
            613
            614
            615
            616
            617
            618
            619
            620
            621
            622
            623
            624
            625
            626
            627
            628
            629
            630
            631
            632
            633
            634
            635
            636
            637
            638
            639
            640
            641
            642
            643
            644
            645
            646
            647
            648
            649
            650
            651
            652
            653
            654
            655
            656
            657
            658
            659
            660
            661
            662
            663
            664
            665
            666
            667
            668
            669
            670
            671
            672
            673
            674
            675
            676
            677
            678
            679
            680
            681
            682
            683
            684
            685
            686
            687
            688
            689
            690
            691
            692
            693
            694
            695
            696
            697
            698
            699
            700
            701
            702
            703
            704
            705
            706
            707
            708
            709
            710
            711
            712
            713
            714
            715
            716
            717
            718
            719
            720
            721
            722
            723
            724
            725
            726
            727
            728
            729
            730
            731
            732
            733
            734
            735
            736
            737
            738
            739
            740
            741
            742
            743
            744
            745
            746
            747
            748
            749
            750
            751
            752
            753
            754
            755
            756
            757
            758
            759
            760
            761
            762
            763
            764
            765
            766
            767
            768
            769
            770
            771
            772
            773
            774
            775
            776
            777
            778
            779
            780
            781
            782
            783
            784
            785
            786
            787
            788
            789
            790
            791
            792
            793
            794
            795
            796
            797
            798
            799
            800
            801
            802
            803
            804
            805
            806
            807
            808
            809
            810
            811
            812
            813
            814
            815
            816
            817
            818
            819
            820
            821
            822
            823
            824
            825
            826
            827
            828
            829
            830
            831
            832
            833
            834
            835
            836
            837
            838
            839
            840
            841
            842
            843
            844
            845
            846
            847
            848
            849
            850
            851
            852
            853
            854
            855
            856
            857
            858
            859
            860
            861
            862
            863
            864
            865
            866
            867
            868
            869
            870
            871
            872
            873
            874
            875
            876
            877
            878
            879
            880
            881
            882
            883
            884
            885
            886
            887
            888
            889
            890
            891
            892
            893
            894
            895
            896
            897
            898
            899
            900
            901
            902
            903
            904
            905
            906
            907
            908
            909
            910
            911
            912
            913
            914
            915
            916
            917
            918
            919
            920
            921
            922
            923
            924
            925
            926
            927
            928
            929
            930
            931
            932
            933
            934
            935
            936
            937
            938
            939
            940
            941
            942
            943
            944
            945
            946
            947
            948
            949
            950
            951
            952
            953
            954
            955
            956
            957
            958
            959
            960
            961
            962
            963
            964
            965
            966
            967
            968
            969
            970
            971
            972
            973
            974
            975
            976
            977
            978
            979
            980
            981
            982
            983
            984
            985
            986
            987
            988
            989
            990
            991
            992
            993
            994
            995
            996
            997
            998
            999
            1000
            1001
            1002
            1003
            1004
            1005
            1006
            1007
            1008
            1009
            1010
            1011
            1012
            1013
            1014
            1015
            1016
            1017
            1018
            1019
            1020
            1021
            1022
            1023
            1024
            1025
            1026
            1027
            1028
            1029
            1030
            1031
            1032
            1033
            1034
            1035
            1036
            1037
            1038
            1039
            1040
            1041
            1042
            1043
            1044
            1045
            1046
            1047
            1048
            1049
            1050
            1051
            1052
            1053
            1054
            1055
            1056
            1057
            1058
            1059
            1060
            1061
            1062
            1063
            1064
            1065
            1066
            1067
            1068
            1069
            1070
            1071
            1072
            1073
            1074
            1075
            1076
            1077
            1078
            1079
            1080
            1081
            1082
            1083
            1084
            1085
            1086
            1087
            1088
            1089
            1090
            1091
            1092
            1093
            1094
            1095
            1096
            1097
            1098
            1099
            1100
            1101
            1102
            1103
            1104
            1105
            1106
            1107
            1108
            1109
            1110
            1111
            1112
            1113
            1114
            1115
            1116
            1117
            1118
            1119
            1120
            1121
            1122
            1123
            1124
            1125
            1126
            1127
            1128
            1129
            1130
            1131
            1132
            1133
            1134
            1135
            1136
            1137
            1138
            1139
            1140
            1141
            1142
            1143
            1144
            1145
            1146
            1147
            1148
            1149
            1150
            1151
            1152
            1153
            1154
            1155
            1156
            1157
            1158
            1159
            1160
            1161
            1162
            1163
            1164
            1165
            1166
            1167
            1168
            1169
            1170
            1171
            1172
            1173
            1174
            1175
            1176
            1177
            1178
            1179
            1180
            1181
            1182
            1183
            1184
            1185
            1186
            1187
            1188
            1189
            1190
            1191
            1192
            1193
            1194
            1195
            1196
            1197
            1198
            1199
            1200
            1201
            1202
            1203
            1204
            1205
            1206
            1207
            1208
            1209
            1210
            1211
            1212
            1213
            1214
            1215
            1216
            1217
            1218
            1219
            1220
            1221
            1222
            1223
            1224
            1225
            1226
            1227
            1228
            1229
            1230
            1231
            1232
            1233
            1234
            1235
            1236
            1237
            1238
            1239
            1240
            1241
            1242
            1243
            1244
            1245
            1246
            1247
            1248
            1249
            1250
            1251
            1252
            1253
            1254
            1255
            1256
            1257
            1258
            1259
            1260
            1261
            1262
            1263
            1264
            1265
            1266
            1267
            1268
            1269
            1270
            1271
            1272
            1273
            1274
            1275
            1276
            1277
            1278
            1279
            1280
            1281
            1282
            1283
            1284
            1285
            1286
            1287
            1288
            1289
            1290
            1291
            1292
            1293
            1294
            1295
            1296
            1297
            1298
            1299
            1300
            1301
            1302
            1303
            1304
            1305
            1306
            1307
            1308
            1309
            1310
            1311
            1312
            1313
            1314
            1315
            1316
            1317
            1318
            1319
            1320
            1321
            1322
            1323
            1324
            1325
            1326
            1327
            1328
            1329
            1330
            1331
            1332
            1333
            1334
            1335
            1336
            1337
            1338
            1339
            1340
            1341
            1342
            1343
            1344
            1345
            1346
            1347
            1348
            1349
            1350
            1351
            1352
            1353
            1354
            1355
            1356
            1357
            1358
            1359
            1360
            1361
            1362
            1363
            1364
            1365
            1366
            1367
            1368
            1369
            1370
            1371
            1372
            1373
            1374
            1375
            1376
            1377
            1378
            1379
            1380
            1381
            1382
            1383
            1384
            1385
            1386
            1387
            1388
            1389
            1390
            1391
            1392
            1393
            1394
            1395
            1396
            1397
            1398
            1399
            1400
            1401
            1402
            1403
            1404
            1405
            1406
            1407
            1408
            1409
            1410
            1411
            1412
            1413
            1414
            1415
            1416
            1417
            1418
            1419
            1420
            1421
            1422
            1423
            1424
            1425
            1426
            1427
            1428
            1429
            1430
            1431
            1432
            1433
            1434
            1435
            1436
            1437
            1438
            1439
            1440
            1441
            1442
            1443
            1444
            1445
            1446
            1447
            1448
            1449
            1450
            1451
            1452
            1453
            1454
            1455
            1456
            1457
            1458
            1459
            1460
            1461
            1462
            1463
            1464
            1465
            1466
            1467
            1468
            1469
            1470
            1471
            1472
            1473
            1474
            1475
            1476
            1477
            1478
            1479
            1480
            1481
            1482
            1483
            1484
            1485
            1486
            1487
            1488
            1489
            1490
            1491
            1492
            1493
            1494
            1495
            1496
            1497
            1498
            1499
            1500
            1501
            1502
            1503
            1504
            1505
            1506
            1507
            1508
            1509
            1510
            1511
            1512
            1513
            1514
            1515
            1516
            1517
            1518
            1519
            1520
            1521
            1522
            1523
            1524
            1525
            1526
            1527
            1528
            1529
            1530
            1531
            1532
            1533
            1534
            1535
            1536
            1537
            1538
            1539
            1540
            1541
            1542
            1543
            1544
            1545
            1546
            1547
            1548
            1549
            1550
            1551
            1552
            1553
            1554
            1555
            1556
            1557
            1558
            1559
            1560
            1561
            1562
            1563
            1564
            1565
            1566
            1567
            1568
            1569
            1570
            1571
            1572
            1573
            1574
            1575
            1576
            1577
            1578
            1579
            1580
            1581
            1582
            1583
            1584
            1585
            1586
            1587
            1588
            1589
            1590
            1591
            1592
            1593
            1594
            1595
            1596
            1597
            1598
            1599
            1600
            1601
            1602
            1603
            1604
            1605
            1606
            1607
            1608
            1609
            1610
            1611
            1612
            1613
            1614
            1615
            1616
            1617
            1618
            1619
            1620
            1621
            1622
            1623
            1624
            1625
            1626
            1627
            1628
            1629
            1630
            1631
            1632
            1633
            1634
            1635
            1636
            1637
            1638
            1639
            1640
            1641
            1642
            1643
            1644
            1645
            1646
            1647
            1648
            1649
            1650
            1651
            1652
            1653
            1654
            1655
            1656
            1657
            1658
            1659
            1660
            1661
            1662
            1663
            1664
            1665
            1666
            1667
            1668
            1669
            1670
            1671
            1672
            1673
            1674
            1675
            1676
            1677
            1678
            1679
            1680
            1681
            1682
            1683
            1684
            1685
            1686
            1687
            1688
            1689
            1690
            1691
            1692
            1693
            1694
            1695
            1696
            1697
            1698
            1699
            1700
            1701
            1702
            1703
            1704
            1705
            1706
            1707
            1708
            1709
            1710
            1711
            1712
            1713
            1714
            1715
            1716
            1717
            1718
            1719
            1720
            1721
            1722
            1723
            1724
            1725
            1726
            1727
            1728
            1729
            1730
            1731
            1732
            1733
            1734
            1735
            1736
            1737
            1738
            1739
            1740
            1741
            1742
            1743
            1744
            1745
            1746
            1747
            1748
            1749
            1750
            1751
            1752
            1753
            1754
            1755
            1756
            1757
            1758
            1759
            1760
            1761
            1762
            1763
            1764
            1765
            1766
            1767
            1768
            1769
            1770
            1771
            1772
            1773
            1774
            1775
            1776
            1777
            1778
            1779
            1780
            1781
            1782
            1783
            1784
            1785
            1786
            1787
            1788
            1789
            1790
            1791
            1792
            1793
            1794
            1795
            1796
            1797
            1798
            1799
            1800
            1801
            1802
            1803
            1804
            1805
            
```


Search completed: January 19, 2003, 05:05:33
Run time : 1945.65 secs

Genome version 5.1.4
Copyright (c) 1994 - 2003 Computer Ltd.

us-09-026-459a-32

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

us-09-026-459a-32

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

us-09-026-459a-32

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

us-09-026-459a-32

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

us-09-026-459a-32

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

January 16, 2003, 15:20:22 : Search time 2911.86 seconds
(without alignments)
18469.506 Million cell updates/sec

7	752.4	22.6	1122	14	HM515881
8	725.4	21.8	727	9	AL597911
9	729.4	21.7	882	14	BC222227
10	699.6	21.1	823	12	BC610661
11	698.2	21.0	792	9	AL564757
12	692.2	20.8	767	12	BC573456
13	691.8	20.8	972	13	HM466194
14	678	20.4	835	13	HM453724
15	660	19.9	672	13	BC1868359
16	643.6	19.4	797	10	BE549278
17	642.2	19.3	681	12	BC616219
18	617	18.6	618	13	HM264009
19	614	18.5	694	10	BE082836
20	605.2	18.2	623	10	AW583181
21	605	18.2	605	9	AL599633
22	596.4	17.9	922	12	BC254543
23	580	17.5	580	9	AL599105
24	575.2	17.3	914	9	AA764411
25	566.4	17.0	588	10	AW368234
26	558	16.8	871	9	AA764485
27	535.2	16.0	598	10	AW582967
28	530	15.9	539	9	AL692790
29	528	15.9	522	9	AL125941
30	516.8	15.6	597	12	BE941246
31	513.2	15.4	594	10	AV715533
32	512.2	15.4	734	12	BC114857
33	505.2	15.2	736	13	BC151568
34	493.8	14.9	525	9	AA258255
35	486.4	14.6	584	12	BC149050
36	484.4	14.6	494	10	AW502887
37	481.4	14.5	484	10	BE168095
38	480.2	14.5	659	13	BC182116
39	476.6	14.3	2083	13	BM456728
40	474.4	14.3	734	13	BC1526982
41	459.2	13.8	497	12	BE944784
42	457	13.8	457	9	AL094215
43	457	13.8	732	9	AL598766
44	454.2	13.7	869	12	BE065749
45	447.2	13.5	832	9	AL646038

ALIGNMENTS

RESULT 1	BM450031	1024 bp	miRNA	linear	EST 05 FEB 2002
LOCUS	AGENCODR1_6393544	NIH_MGC_72	Homo sapiens	cdna clone IMAGE:5528047	
DEFINITION	5' mRNA sequence				
ACCESSION	BM450031				
VERSION	BM450031.1	GI:18496071			
KEYWORDS	EST				
SOURCE	human				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo				
AUTHORS	1 (bases 1 to 1023)				
TITLE	NHL-MGC http://marc.ncbi.nlm.nih.gov/				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D., email: rstraus@nhi.nih.gov Tissue Procurement: ATCC/DCT10/DTF cdna Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L1AM1204 Row: d Column: 22 High quality sequence stop: 643 Location/Qualifiers 1..1024				

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	910.4	27.4	1024	BM450031	BM450031 AGENCODR1
2	909.4	27.4	1024	BM561083	BM561083 AGENCODR1
3	896.6	25.8	1091	BC439221	BC439221 AGENCODR1
4	844.8	23.4	881	BC222227	BC222227 AGENCODR1
5	824.6	23.4	1364	AK011246	AK011246 Mus muscu
6	786	23.7	1001	BM475603	BM475603 AGENCODR1

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_72"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/vector="Orikan skin Vector pMV-Spork16; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primers: oligo d1. Average insert size 2 kb. Library constructed by Life Technologies."
BASE COUNT      44 a      211 c      197 g      281 t
ORIGIN

Query Match      27.4%; Score 910.2; DB 13; Length 1023;
Best local Similarity 95.2%; Prod. No. 3.2e-165;
Matches 972; Conservative 0; Mismatches 43; Indels 6; Gaps 3;

QY 1911 CAACGCAAGCAATACAGACCTTAATCAAAATCAATGTAACAGCAATACAGCAATCTCC 1970
DB 1 CAACGCAAGCAATACAGACCTTAATCAAAATCAATGTAACAGCAATACAGCAATCTCC 60
QY 1971 TCATCTGTCTGAGGAGACATTCNAACGTCTTTTCATCAAAAGANANANATATCAATATAT 2040
DB 63 TCACTGCTGTGAGGAGACATTCNAACGTCTTTTCATCAAAAGANANANATATCAATATAT 120
QY 2041 TATAGTATCTTAACTGCGGCTTCATGCGACACATCAAAACAAATATTTTGCAGTATCC 2090
DB 121 TATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 180
QY 2091 TCTACATGAGCGCGCTTACCTGCTGACATCAATACCTGACATTCCTGCAAGCGCTTACAGCTT 2150
DB 181 TCCACCAAGCGCGCTTACCTGCTGACATCAATACCTGACATTCCTGCAAGCGCTTACAGCTT 240
QY 2151 TCTACATGAGCGCGCTTACCTGCTGACATCAATACCTGACATTCCTGCAAGCGCTTACAGCTT 2210
DB 241 TCTACATGAGCGCGCTTACCTGCTGACATCAATACCTGACATTCCTGCAAGCGCTTACAGCTT 300
QY 2211 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2270
DB 401 ATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 360
QY 2271 AGTATCAATATGCTGCAATATATGCTGCAATATATGCTGCAATATATGCTGCAATATATGCT 2330
DB 461 AGTATCAATATGCTGCAATATATGCTGCAATATATGCTGCAATATATGCTGCAATATATGCT 420
QY 2441 ATGTAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2490
DB 421 ATGTAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 2491 GAAAAAATACCTTTGATATTTGAGGATCAGATGAGGACATGGAAGTAAACATCTCC 2450
DB 481 GAAAAAATACCTTTGATATTTGAGGATCAGATGAGGACATGGAAGTAAACATCTCC 540
QY 2451 AGGACATCTCAAAATTTACGACAACTCTGATGATGATGATGATGATGATGATGATGATGAT 2510
DB 541 AGGACATCTCAAAATTTACGACAACTCTGATGATGATGATGATGATGATGATGATGATGAT 600
QY 2511 AAATGCAAAATATGATATATATGATATATATATATATATATATATATATATATATATAT 2570
DB 601 AAATGCAAAATATGATATATATGATATATATATATATATATATATATATATATATATAT 660
QY 2571 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2630
DB 661 GAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 2641 TAACTTCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2690
DB 721 TAACTTCCCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 2691 AAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2748
DB 781 AAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH_MGC_67"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/vector="Orikan skin Vector pMV-Spork16; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primers: oligo d1. Average insert size 1.75 kb. Library constructed by Life Technologies."
BASE COUNT      330 a      221 c      197 g      299 t
ORIGIN

Query Match      27.4%; Score 909.4; DB 13; Length 1051;
Best local Similarity 98.4%; Prod. No. 4.6e-165;
Matches 938; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1069 AAAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1128
DB 1 AAAGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 1129 TTCTATTACCGAGTAATGGAATCCATGCTTAAATATCAACAGCAAGCAATATCCATTCAA 1188
DB 61 TTCTATTACCGAGTAATGGAATCCATGCTTAAATATCAACAGCAAGCAATATCCATTCAA 120
QY 1189 AATTTACCAACATCTCAATGACACACATTTTTCATATGCTTTTATGCGTCCCTCTCT 1248
DB 1189 AATTTACCAACATCTCAATGACACACATTTTTCATATGCTTTTATGCGTCCCTCTCT 1300

```



```

27 604 CTCTGAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 664
36 605 CTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 420
37 606 AATCTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 724
38 607 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 480
39 608 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 784
40 609 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 540
41 610 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 844
42 611 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 600
43 612 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 904
44 613 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 656
45 614 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 962
46 615 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 720
47 616 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 989
48 617 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT 747
49 618 TTTTCTGTAAGAAATGAAATTAATATATATAGAGAGTGAAGAAATGTTTATTTTCAAAAAATTTT

RESULT 10
LOCUS 10661
DEFINITION 626118481 NIH_MGC_50 Homo sapiens cDNA clone IMAGE:4737634.57.
VERSION 106610661
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Graptata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 823)
AUTHORS NIH MGC http://www.ncbi.nlm.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstraus@nhi.nih.gov
Tissue Procurement: BMD/2TP
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: LCM1605, row: 0, column: 11
High quality sequence stop: 709.
FEATURES
Location/Qualifiers
1..823
Zonedism "Homo sapiens"
Zdb_xref: "taxon:9606"
Zclone: "IMAGE:4737634"
Zclone_lib: "NIH_MGC_50"
Ztissue_type: "adenocarcinoma"
Zlab_host: "pH108 (11 phage resistant)"
Zlab_origin: "prostate; Vector: pBK11H (Clontech);
Site 1: S11 (sequence); Site 2: S11 (sequence)
Site 3: Double stranded cDNA was prepared from cell line RNA.
5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATATGAC-3' and 3' adaptor
sequence: 5'-ATCTACAGCGGACGCGGCGGACAG-dT(30)NN-3'
(Insert: A, C, G or N - A, C, G, or T). Average
insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies
contained inserts by PCR. This library was enriched for

```

```

full length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library.
BASE COUNT 262 a 148 c 142 g 271 t
ORIGIN
Query Match 21.1% Score 699.6; DB 12; Length 823;
Best Local Similarity 94.8%; Pred. No. 1, 2e 124;
Matches 779; Conservative 0; Mismatches 44; Indels 9; Gaps 0;
QY 2451 AGGACAGCAGCAAAATTCAGTACACAAAGTGTAGAAATGCTTCTACCTGCAATATGAAATGA 2510
DB 1 AGGACAGTGTCAAAATTCAGTACAAAGTGTAGAAATGCTTCTACCTGCAATATGAAATGA 60
QY 2511 AAAGTACAGAAAATGAAATGATAGATATATATATATATATATATATATATATATATATAT 2570
DB 61 AAAGTACAGAAAATGAAATGATAGATATATATATATATATATATATATATATATATATAT 120
QY 2571 GAGCTTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2640
DB 121 GAGCTTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
QY 2641 TAACTTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2690
DB 181 TAACTTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 240
QY 2691 AAAATGTCAGATGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2760
DB 241 AAAATGTCAGATGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 400
QY 2761 TATTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2810
DB 301 TATTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 460
QY 2811 GTTCTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2870
DB 361 GTTCTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 2871 GAGTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2940
DB 421 GAGTGTGTGAAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
QY 2941 TATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2990
DB 481 TATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 2991 TACAAATGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3060
DB 541 TACAAATGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
QY 3061 TATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3107
DB 601 TATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
QY 3108 ATTTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3167
DB 661 ATTTTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 717
QY 3168 AAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3227
DB 718 AAAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 776
QY 3228 TACTTATCAAAACATATATATATATATATATATATATATATATATATATATATATATATAT 3269
DB 777 TACTTATCAAAACATATATATATATATATATATATATATATATATATATATATATATATAT 816
RESULT 11
LOCUS AL563757
DEFINITION AL563757 L11_NF1301_NHC4 Homo sapiens cDNA clone F300000/AY05.4
prime, mRNA sequence.
ACCESSION AL563757

```




cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Amersham Bioscience Corporation
 clone distribution: MAC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:

FEATURES SOURCE

Technologies.			
BASE COUNT	330 d	221 c	197 d
ORIGIN			
Alignment Scores:			
Prod. No.:	1,550-154		Length:
Score:	151.00		Matches:
Percent Similarity:	93.25%		6
Best Local Similarity:	91.50%		Mismatches:
Query Match:	45.48%		18
DB:	13		Indels:
			5
			Gaps:
			1
			1051
			412
			Others

US-09-026-459A-33 (1-851) x HM561083 (1 1051)

356	AlaValGlyGlnGlyCysValGluIleGlySerGluArqTyrLeuGlyValArqLeu	475	QY
4	GTCCTGGAGAGGCTCTCTGCAAAATTCGATACAGAGGATACAAATCTGGAGTCTGCTTGG	63	DB
376	TyrTyrArqValMetGluSerMetLeuLysSerGluGluGluArqLeuSerIleGluAsn	395	QY
64	TATTACGAGTAATAGGAAATCGATGCTTAATACAGAAAGAAAGATATGATTAAAAAT	123	DB
396	PheSerLysLeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaLysAlaLeuGlu	415	QY
124	TTTAGCAAACTCTGTAAACACAAACATTTTCAATAGCTCTATTCGGAGTCTCTCTTCAG	183	DB
416	ValValMetAlaThrTyrSerArqSerThrSerGluAsnLeuAspSerGlyThrAspLeu	435	QY
184	GTTGTAAATGGCCATATATAGCAAGATCATCTCAGAACTCTGATTTGGAGAAATAATTTG	243	DB
436	SerPheProTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIle	455	QY
244	TCTTCCCATGGATCTCGAATGGCTTAAATTTAANAATTTTGATTTTAAAGTGATC	403	DB
456	GluSerPheIleLysAlaGluLysAsnLeuThrArqGluMetIleLysHisLeuGluArq	475	QY
304	GAAGACTTTTATCAAGAGCAAAAGCGAATTCGAAAGCAAGAAATGATAAAATATATAGAACCA	463	DB
476	CysGluHisArqIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeu	495	QY
484	TGTGAATCATGAAATCATGAATGCTTGTATGGCTCTCAATTCATCATTTATTATGATCT	423	DB
496	IleCysGlnSerLysAspArgGluGlyProThrAspHisLeuLeuSerAlaLysProLeu	515	QY
424	ATTAAATCAATCAAGGATCGAGAAAGCAACATGATATACCTTGAAATCTCTCTCTCTCT	483	DB
516	AsnLeuProLeuGluAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArqSer	535	QY
484	AATCTTCCTCCCAAGATAATACACATGACAGCAGCAGATGATATCTTCTCTCTTCAAGATCT	543	DB
536	ProLysLysLysGlySerThrThrArqValAsnSerThrAlaAsnAlaGluThrGluAla	555	QY
544	CGAAGAGAAAAGGTTCAATATGAGGTGTAAATTTGATATGAAAGCAAGCAATAGCA	603	DB
556	ThrSerAlaPheGlnThrGlnLysProLysLysSerThrSerLeuSerLeuPheLys	575	QY

607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000
607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722																																																																																																																																																																																																																																																																																						

BM561083	BM561083	1051 bp	linear	EST 20 FEB 2002
LOWIS	LOWIS		mRNA	
DEFINITION	AF096087.6:661-144 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550965			
			5' cDNA sequence	

A. FESSLIN $\mathbb{M}^2, 1084$

Accession	Accession
VRK510N	HM5510
KFYW0K15	HS151

$$M = \mathbb{Z}_2$$

MSINVM
Homo sapiens
Fukuyama, Ma

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

DIFFERENCE
AUTHORS
(bases 1 to 105)
NIM MCG DIFFERENCE

Public National Institutes of Health, Md.

BOOKS. Unpublished (1999)
Contact: Robert Str

Email: craybes@gmail.com

ISSUE PROCUREMENT: ALCC

THE UNIVERSITY OF CHICAGO

```

DB 604 AACCTAGGCTTCCACACACACAGCCATCAAAATACCTCTCTTCACAGTGTATATAAA 663
QY 676 LysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuCysGluArgLeuLeuSerGlu 595
DB 604 AAATGCTAAGCTAGCTATGCTGGCTAAATACACTTCTGAAAGCTCTGCTCTGAC 723
QY 676 HisProGluGluGluHisLeuLeuTyrPheArgPheGluHisThrLeuLeuAsnGlyPhe 615
DB 724 CACCGAATATAGACATATCATCTGACAGCTTTTCAGACACACGCGGACAGAAAGATAT 783
QY 676 GluLeuGluTyrAspArgHisLeuAspGlnLeuMetMetCysSerMetTyrGlyLeuCys 635
DB 784 CAACATATACACACACACACACACACACACACACACACACACACACACACACACAC 843
QY 676 LysValTyrAspLeuAspLeuLysPheCysLeuLeuValTyrLeuValTyrLeuValTyr 655
DB 844 AAAGTGAAGATATAGACATATCAATTAATTTCAAAATCTGTGTAAACAGATACAGATG 903
QY 676 HisAlaValGlnGluThrProGlyArgValLeuLeuLeuGlyGlnGluThrTyrAspSer 674
DB 903 TCACTGTCNTGACGACATTCATCAACCTGTGTGTGATGCAAGAAAGAGAGATATATATG 963
QY 676 IleIleVal-----PheTyrAspSerValIleMetMetGluArgLeuGlyPheAsp 691
DB 964 TATAATACGATATGCTATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

RESULT 4
LOCUS BQ220275 881 bp mRNA linear EST 02-MAY-2002
DEFINITION AF090007.7572421 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575683
5', mRNA sequence.
ACCESSION BQ220275
VERSION BQ220275.1 GI:20401675
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
EXTRACT: Mammalia; Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 881)
AUTHORS NIH-MGC http://mgi.nhl.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contract: Robert Strausberg, Ph.D.
Email: ccaps-rsmall.nhl.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (IIN)
DNA Sequencing by: Amersham Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLML3287 row: e column: 03
High quality sequence stop: 690.
FEATURES
location/Qualifiers
1..881
/organism="Homo sapiens"
/clone="IMAGE:5575683"
/clone="IMAGE:5575683"
/clone="IMAGE:5575683"
/tissue_type="embryonal carcinoma, cell line"
/label="organ: testis; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; cloned indirectly; oligo dT primed
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies
Note: this is a NIH_MGC library."
BASE 1..881 313 a 149 c 155 g 264 t
ORIGIN
Alignment Scores: 1.17e-143 Length: 881
Pred. NO.: 1448.00 Matches: 288

```

```

US-09-026-459a-33 (1-881) x BQ220275 (1-881)
Percent Similarity: 98.30% Conservative: 1
Best Local Similarity: 97.96% Mismatches: 3
Query Match: 33.12% Indels: 2
DB: 14 Gaps: 0

QY 131 MetCysAspSerLeuValTyrSerPheGlnMetCysValLeuAspTyrPhe 150
DB 3 ATGAAGAATGAATGAGGATATTTATTCAGTATATGCTGTGCTGTGACTATTTTAT 62
QY 151 CysLeuSerTyrPheMetLeuLeuGlySerLeuValTyrLeuValTyrLeuValTyr 170
DB 63 AAATCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 122
QY 171 GlySerTyrArgThrPheArgThrGlnGlyGluAsnArgLeuAlaArgLeuAlaLys 190
DB 123 GCTTCACCTTGAAACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 182
QY 191 GluAsnAspThrArgTyrLeuGluValLeuCysTyrSerHisGlnCysAspHisCasp 210
DB 183 GAAATCATATCAAGATATTTTCAAGTTCTCTATTAATCAATATCAATATCAATAT 242
QY 211 ValGlyAsnValTyrPheLeuValTyrPheLeuValTyrPheLeuValTyrPheLeu 230
DB 243 GTCAAAATGTTTATTTCAAAAATTTTATACCTTTTATCAATTTCTGCTGCTGCTGCT 302
QY 231 SerAspGlyLeuProGluValGluAsnLeuSerLeuArgTyrGluGluLeuTyrLeu 250
DB 303 TCAATGAGCTTTCAGAGGCTTGAAATCTTCTAAACGATACGAAACAATTTATCT 362
QY 251 AsnLysAspLeuAspAlaArgGlnPheLeuAspHisAspLysThrLeuGlnThrAsp 270
DB 363 AATAAATCATATATATCAAGATATTTTCAATATTTTCAATATTTTCAATATTTT 422
QY 271 IleAspSerPheGluThrGlnArgThrProArgLysSerAsnLeuAspGluLeuVal 290
DB 423 ATACACACTTTCACACACACACACACACACACACACACACACACACACACACAC 482
QY 291 ValIleProPheHisThrProValArgThrValMetLeuValTyrLeuGlnLeuMet 310
DB 483 GTAAATCTTCCACACACACACACACACACACACACACACACACACACACACAC 342
QY 311 IleLeuAsnSerAlaSerAspGlnProSerGlnAsnLeuLeuSerTyrPheAsnAsn 330
DB 543 ATTTTAAATTCACCAATATGATCAATTTTCAAAAATCTGATTTTATTTTATTTT 402
QY 331 ThrValAspThrGlnSerLeuLeuValTyrPheLeuValTyrPheLeuValTyr 350
DB 603 ACATGTAATCTCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 462
QY 461 HisLysPheAlaLysAlaValTyrGlnTyrCysValTyrLeuGlySerGlnArgTyr 370
DB 663 GAGAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT 422
QY 371 LeuGlyValArgLeuTyrTyrArgValMetGluSerMetCysLysSerGluLeuGlu 390
DB 723 GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 782
QY 391 LeuSerLeuGlnAsnPheSerLysLeuValTyrPheLeuValTyrPheLeuVal 410
DB 783 TTAATCATTTCAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 842
QY 410 GluLeuValGlnGluValValMetLeuValTyrTyrSer 422
DB 843 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880

RESULT 4
BM475603
LOCUS BM475603 1001 bp mRNA linear EST 05-FEB-2002
DEFINITION AF090007.7572421 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575683
5', mRNA sequence.
ACCESSION BM475603

```



```

14 241 GAAATATATAAAACAGTGGTATTAACGATTAATGGTTCCAGCTGAGAACACGAGAGAGGT 400
15
16 180 GAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 194
17
18 301 CAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 360
19
20 200 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 219
21
22 361 CAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 420
23
24 220 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 239
25
26 421 ATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 480
27
28 240 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 259
29
30 481 CAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 540
31
32 260 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 279
33
34 541 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 600
35
36 280 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 299
37
38 601 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 660
39
40 300 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 319
41
42 661 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 720
43
44 319 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 338
45
46 721 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 780
47
48 338 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 353
49
50 781 LAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 839
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

```

```

/Note: *organ: skin; Vector: pCMV Sp6/6; Site: 1; Not 1;
Site: 2; Salt: Cloned unidirectionally. Primer: 1000 bp.
Average insert size 2 kb. Library constructed by Life
technologies.
BASE COUNT 254 a 210 c 137 g 244 t 1 others
ORIGIN
Alignment Scores:
Prod. No.: 4.4e-118 Length: 846
Score: 1206.00 Matches: 244
Percent Similarity: 92.91% Conserved: 5
Best Local Similarity: 91.04% Mismatches: 12
Query Match: 27.63% Indels: 7
DB: 14 Gaps: 5
US-09-026-459a-33 (1-851) x HM453724 (1-845)
QY 473 LeuGluArgCysGlyHisArgLleMetGluSerLeuAlaTrpLeuSerAspSerProLeu 492
Db 8 GTCCGACGATGTGAATATCGAATCATGAAATCGCTTGATGCTCTAGATTAATTTA 67
QY 493 pheAspLeuIleLysGlnSerLysAspArgLleGlyProThrAspHisLeuGlnSerAla 512
Db 68 TTGATCTTATTAAATCAATCAAAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 127
QY 513 CysProLeuAsnLeuProLeuGlnAsnHisThrAlaAlaAspMetTyrLeuSerPro 532
Db 128 TGTCTCTTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 187
QY 533 ValArgSerProLysLysLysLysSerThrThrArgValAsnSerThrAlaAsnAlaGln 552
Db 188 GTAGATCTCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 247
QY 553 ThrGlnAlaThrSerAlaPheGlnThrGlnLysProLysSerThrSerLeuSerLeu 572
Db 248 AACAAAGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA 607
QY 573 PheTyrLysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuCysHisAlaLeu 592
Db 308 TTTTAAATAAAAGCTGATCGGCTAGCCATATCTGCGCTAAATCAAAATCAAAATCA 667
QY 593 LeuSerGluHisProLysLeuGlnHisLleThrThrLeuPheLeuHisThrLeuHis 612
Db 368 CTCTCTGACACACACACACACACACACACACACACACACACACACACACACACACAC 427
QY 613 AsnGlnTyrGlnLeuMetArgAspArgHisLeuAspGlnLleMetMetCysSerMetTyr 632
Db 428 AATGAGTATGAATCATCATGAGAGACAGGATTTGACCAAAATATATATATATATATAT 487
QY 633 GlyIleCysLysValLysAsnLleAspLeuLysPheLysLleLleValThrAlaTyrLys 652
Db 488 GGCATATGCAAGCTCAAGATATAGACCTTAAATTCAAAATCAATCAAAATCAAAATCA 547
QY 653 AspLeuProHisAlaValGlnGlnThrPheLysArgValLeuLleLysGlnLleLys 672
Db 548 GATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 607
QY 673 AspSerLleLleValPheLysAsnSerValPheMetGlnArgGlnLysThrAsnLeu 692
Db 608 GATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 667
QY 693 GlnTyrAlaSerThrArgProProProProProProProProProProProProPro 711
Db 668 CAGTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 727
QY 711 rProTyrLysProProProProProProProProProProProProProProProProPro 726
Db 728 GAAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 787
QY 727 LleTyrLleSerProLysLys 733
Db 788 TATATTTTCAACGCTGGA 809

```

FEATURES

Source:
 1. 845
 2. 845
 3. 845
 4. 845
 5. 845
 6. 845
 7. 845
 8. 845
 9. 845
 10. 845
 11. 845
 12. 845
 13. 845
 14. 845
 15. 845
 16. 845
 17. 845
 18. 845
 19. 845
 20. 845
 21. 845
 22. 845
 23. 845
 24. 845
 25. 845
 26. 845
 27. 845
 28. 845
 29. 845
 30. 845
 31. 845
 32. 845
 33. 845
 34. 845
 35. 845
 36. 845
 37. 845
 38. 845
 39. 845
 40. 845
 41. 845
 42. 845
 43. 845
 44. 845
 45. 845
 46. 845
 47. 845
 48. 845
 49. 845
 50. 845
 51. 845
 52. 845
 53. 845
 54. 845
 55. 845
 56. 845
 57. 845
 58. 845
 59. 845
 60. 845
 61. 845
 62. 845
 63. 845
 64. 845
 65. 845
 66. 845
 67. 845
 68. 845
 69. 845
 70. 845
 71. 845
 72. 845
 73. 845
 74. 845
 75. 845
 76. 845
 77. 845
 78. 845
 79. 845
 80. 845
 81. 845
 82. 845
 83. 845
 84. 845
 85. 845
 86. 845
 87. 845
 88. 845
 89. 845
 90. 845
 91. 845
 92. 845
 93. 845
 94. 845
 95. 845
 96. 845
 97. 845
 98. 845
 99. 845
 100. 845

FEATURES

Source:
 1. 845
 2. 845
 3. 845
 4. 845
 5. 845
 6. 845
 7. 845
 8. 845
 9. 845
 10. 845
 11. 845
 12. 845
 13. 845
 14. 845
 15. 845
 16. 845
 17. 845
 18. 845
 19. 845
 20. 845
 21. 845
 22. 845
 23. 845
 24. 845
 25. 845
 26. 845
 27. 845
 28. 845
 29. 845
 30. 845
 31. 845
 32. 845
 33. 845
 34. 845
 35. 845
 36. 845
 37. 845
 38. 845
 39. 845
 40. 845
 41. 845
 42. 845
 43. 845
 44. 845
 45. 845
 46. 845
 47. 845
 48. 845
 49. 845
 50. 845
 51. 845
 52. 845
 53. 845
 54. 845
 55. 845
 56. 845
 57. 845
 58. 845
 59. 845
 60. 845
 61. 845
 62. 845
 63. 845
 64. 845
 65. 845
 66. 845
 67. 845
 68. 845
 69. 845
 70. 845
 71. 845
 72. 845
 73. 845
 74. 845
 75. 845
 76. 845
 77. 845
 78. 845
 79. 845
 80. 845
 81. 845
 82. 845
 83. 845
 84. 845
 85. 845
 86. 845
 87. 845
 88. 845
 89. 845
 90. 845
 91. 845
 92. 845
 93. 845
 94. 845
 95. 845
 96. 845
 97. 845
 98. 845
 99. 845
 100. 845

was constructed by Benito Soares and M. Fatima Bonaldo.
 BASE COUNT 251 a 178 c 217 g 261 t 1 others
 ORIGIN

Alignment Scores:
 Prod. No.: 7,496-192 Length: 914
 Score: 1056.50 Matches: 244
 Percent Similarity: 82.14% Conservative: 19
 Best Local Similarity: 75.97% Mismatches: 53
 Query Match: 24.17% Indels: 6
 Gaps: 2

US 09 026 459A 33 (1-851) x AA764411 (1-914)

QY 259 PheLeuAspHisAspLysThrLeuGluThrAspSerPheGluThrGlnArg 278
 DB 914 TTTTGGATCGAATAAAGGCGGAGTGT TTTTAAAGGCGGAG 857
 QY 276 ThrProArgLysSerAsnLeuAspGluValAsnValIleGlnProHisThrProVal 298
 DB 856 AGCGAGCGAAGAAAGGCTTGT GAAGCGGCGAAGCGGAGTTACTCGCGCGACGCGCGGTT 798
 QY 299 ArgThrValMetAsnThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGln 318
 DB 797 TGGTGTGTTTGGAACTATCGAACAATTTATGTCGATTTTAAATCGCGCAATGACCGC 738
 QY 319 ProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIle 338
 DB 737 GCGTGAATATGCTTTTGGTGTCTAATATTCGCGCGGATGAAAGGAAATTCGC 679
 QY 339 LeuLysAsnValLysAspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValGly 358
 DB 678 CTAAAGTGTGT AAGAGGTTTGGGATATGCTTAAAGAGAGGTTCTTA---ACGCTGTGG 623
 QY 359 GlnHisCysValGlnIleLysSerIleArgTyrLysLeuGlyValArgLeuTyrTyrArg 378
 DB 622 CTAAGCTGTGTGACATGAGACACAGGATATAAACTTCGAGTCGACATGCTATACCGT 563
 QY 379 ValMetGluSerMetLeuLysSerGlnGluGluArgLeuSerIleLeuAsnPheSerLys 398
 DB 562 GTGATGAAATGATGCTTATGATACAGAGAGAGAGGTTTGTCATTCGAAATTTAGCAA 503
 QY 399 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaLysAlaLeuGluValMet 418
 DB 502 CTCTTAATATGAAATATTTTCATATGCTTTACTGCTGTGTGTCTTGAGTTGTAATG 443
 QY 419 AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro 438
 DB 442 GTTAAATAGTACAGACAGTACATTCGACGACATTCGATTCGCAACAGATTCGCTTCGCG 383
 QY 439 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPhe 458
 DB 382 TGTATTCGAAAGGTTACTTAATTTAAAGGCTTTGATTTTACAAAGGTTGTAAGGTTT 323
 QY 459 ThrLysAlaGlnLysAsnLeuThrArgGluMetIleLysHisLeuGluArgCysGluHis 478
 DB 322 ATCAAGTGGAGAGCAATTTGAAAGAGATGATAAATATTAGAAAGATGTGAGCAT 263
 QY 479 ArgIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysGln 498
 DB 262 CCAATCAAGCAAGCTTCGATGAGCTTCAGATTCAGCTTTATTTGATCATCATTAAGCAG 203
 QY 499 SerLysAspArgGlnGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuPro 518
 DB 202 TCAGAGATGGAAGAGAGCT---GATAAGCTTGAAGCTGCTTGTCTCTGAGCTGCT 146
 QY 519 LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 538
 DB 145 CTCAAGGTTAAATATATGAGATATATCTTTCTGCTGCTAAGAGATCTCAAGAGAA 86
 QY 539 LysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 558
 DB 86 AGAAATTCACATATAGTGTAAATTTCTGCTGCAAAATACAGACACACAGGCTCAGCC 26

QY 559 PheGlnThrGlnLysProLeuLys 566
 DB 25 TTTCTACTCAGAGGATTTGAAA 2

RESULT 10
 AL563757/c

LOCUS AL563757 LTL.NF1.001.NB4 792 bp mRNA Human EST 16-FEB 2001
 DEFINITION AL563757 LTL.NF1.001.NB4 Homo sapiens cDNA clone CS040007YA05, a
 prime, mRNA sequence.

ACCESSION AL563757
 VERSION AL563757.1 GI:1291464
 KEYWORDS EST.
 SOURCE human.

ORGANISM
 Homo sapiens

REFERENCE
 AUTHORS Li, W.-H., Gruber, C., Jessup, J., and Polayes, D.
 TITLE Full length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope

Genoscope - Centre National de Sequençage
 BP 191 91006 Evry cedex - France
 Email: seq@genoscope.cns.fr, Web: www.genoscope.cns.fr,
 location/Qualifiers

1. 792
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS000007YA05"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /vector="pJen4, Brain, Vector, pMSP6ki 6, 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double stranded cDNA was digested with NotI and
 cloned into the NotI and EcoRV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed
 by Life Technologies. Contact: Feng Liang Life
 Technologies, a division of Invitrogen 9800 Medical Center
 Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
 8371 Email: fliang@lifetech.com URL:
 http://fulllength.invitrogen.com"

FEATURES
 source

BASE COUNT 246 a 148 c 117 g 255 t 26 others
 ORIGIN

Alignment Scores:

Prod. No.: 1,450-100 Length: 792
 Score: 1043.00 Matches: 217
 Percent Similarity: 84.53% Conservative: 7
 Best Local Similarity: 81.89% Mismatches: 48
 Query Match: 23.86% Indels: 4
 Gaps: 0

US-09-026-459A-33 (1-851) x AL563757 (1-792)

QY 163 LysThrAlaValIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArg 182

DB 791 AAACAGCTKTATACCGATTAATGGTTTCACCTGCAATACACAGGTCAGTTCAGAAACAG 742

QY 183 SerAlaArgIleAlaLysGlnLeuGluAsnAspThrArgIleIleGluValLeuLysLys 202

DB 731 ACTGACGGGTAGGACACCAATTAATTTGATTAATTAATTTGAGTTCTTGTAAA 672

QY 203 GluHisGluCysAsnIleAlaGluValLysAsnValTyrPheCysAspPheIleProPhe 222

DB 673 CAACATGAATGATATATACGTCAGCTGCAAAAGCTTTTWTTCAGAAATTTTATACCTTT 612

QY 223 MetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGluAsnLeuSerLys 242

DB 611 ATCAATCTCTTGGACTTGTACATATWATGCACTTACAGGTTGAAATCTCTTCTAAA 592


```

10 564 GTGTGATGATATCATGTTTAAATCAGAGAAAGAAAGTTTGTCATTCAGAAATTTTACCA 505
11 598 yslleuLeuAsnAspSerLeuPheHisMetSerLeuLeuAlaCysAlaLeuGluValValM 418
12 604 CACCTTAAATCAGACAACTTTCATATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 445
13 618 etAlaThrTyrSerArgSerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPheP 438
14 644 TGGTATAGTATAGCAGACAACTTACATGCACTTCCTGATCTGCAACACATTCCTGCTGCC 385
15 648 ttrpLeuLeuAsnValLeuAsnLeuGlySerLeuPheAspPheLeuValLeuGluSerP 458
16 684 GGTGGATCTGACAGCTATTTATTTTAAAGCTTTGATTTTACAAATGATTTGAAAGTT 325
17 698 heLeuLysAlaLeuThrLysLeuThrArgLeuMetLeuLysHisLeuGluArgCysGluH 478
18 724 TTAATTAAGTCCAAATTAATTAACAAACAAATCATAAAATATTAACAACATGCTGACG 265
19 748 isArgLeuMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuLeuLysG 498
20 764 ATCAATATCAAGAACTGCTGCAAGCTTTCAGATTCACCTTATTTGATCTGATTAAGC 205
21 798 InsLeuLysAspArgGluLeuProThrAspHisLeuGluSerAlaCysProLeuAsnLeuP 518
22 804 ATCAATATCAAGAACTGCTGCAAGCTTTCAGATTCACCTTATTTGATCTGATTAAGC 148
23 818 trolLeuLysAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysL 538
24 848 AAATCAATCTGCTGCAAGCTTTCAGATTCACCTTATTTGATCTGATTAAGC 88
25 858 yslLysGlySerThrThrArgValAsnSerThrAlaAsnAlaLeuThrGluAlaThrSerA 558
26 878 AAATCAATCTGCTGCAAGCTTTCAGATTCACCTTATTTGATCTGATTAAGC 28
27 898 labeledInThrLysProLeuLys 566
28 908 CATTGATATTAAGAACTGCTGCAAGCTTTCAGATTCACCTTATTTGATCTGATTAAGC 2

```

RESULT 15
 6025454
 DEFINITION 6025454 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4472611 5',
 mRNA sequence.
 ACCESSION 6025454
 VERSION 6025454.1 GI:12763359
 KEYWORDS EST.
 SOURCE Human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 922)
 NIH_MGC http://map.ncbi.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: rstraus@nigms.nih.gov
 Tissue procurement: AFCC
 cDNA library preparation: Life Technologies, Inc.
 cDNA library arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA sequencing by: Invitrogen, Inc.
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://imgc.lnc.nih.gov
 Plated: LNL0293 row: 1 column: 20
 high quality sequence stop: 643.
 Location/Qualifiers
 1..922
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4472611"
 /clone_lib="NIH_MGC_90"
 /issue_type="cDNA library", cell line"
 /lab_host="DH10B (phage-resistant)"

/note="origin: liver; Vector: pMV-Sp646; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; clones of primed
 Average insert size 1.7 kb. Library enriched for
 full length clones and constructed by life technologies.
 Note: this is a NIH_MGC library."

BASE COUNT 307 a 214 c 204 g 197 t
 ORIGIN
 Alignment Scores:
 Prod. No.: 40-96 Length: 922
 Score: 1004.00 Matches: 224
 Percent Similarity: 90.51% Conservatives: 5
 Best Local Similarity: 88.54% Mismatches: 21
 Query Match: 22.96% Indels: 8
 DB: 12 Gaps: 0
 02-09-026 459a 33 (1 853) X 6025454 (3 922)
 QY 602 IleIleTrpThrLeuPheGlnHisThrLeuGluAsnGlu TyrGluLeuMetArgAspAr 621
 DB 1 ATCATCTGGACCTTTTCAGCAGCAACCTTCACAAAGAGTAAAGAACTATCAGACAGATAG 60
 QY 621 qHisLeuAspGlnIleuMetMetCysSerMetTyrGlyIleCysLysValLysAsnLys 641
 DB 61 GATTCGACAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119
 QY 641 pLeuLysPheCysIleIleValThrAlaTyrLysAspLeuProHisAlaValIleGluTh 661
 DB 120 CCTTAAATTCAAAATGAT-GTAAACAGCAATAAAGCATCTTCCTGATGATGATGATGATGAT 178
 QY 661 rPheLysAsnValLeuIleLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 681
 DB 174 ATTCAAACCTGCT-TTGATGCAAAACAGCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 237
 QY 681 rValPheMetGlnArgLeuLysThrAsnIleLeuGlnIleValThrAlaSerThrArgProThr 701
 DB 238 GGTCTTATGTCAGACAGTCAAAACAAATAT-TTGATGATGATGATGATGATGATGATGATGAT 296
 QY 701 rLeuSerProIleProHisIleProArgSerProTyrLysPheProSerProSerProSer 721
 DB 297 CTCTTCACCAATACCTCAGATCTTCCTGGAAGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 356
 QY 721 qIleProGlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIleSerGluGln 741
 DB 357 GATTCCTGGAGGAACTATATATTCATGCTGCAAGCTTCAAGTAAAGTAAAGTAAAGTAAAG 416
 QY 741 rLeuProThrProThrLysMetThrProArgSerArgIleLeuValSerIleGlyLysGln 761
 DB 417 TCTGCCAC 476
 QY 761 rPheGlyThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgVal 781
 DB 477 ATTCGGCAGCTTCAGAACTTCAGAAAAATAAATACATGATGATGATGATGATGATGATGATGAT 536
 QY 781 rLeuLysArgSerAlaGlnGlySerAsnProLysProLeuLysLysSerArgPheAs 801
 DB 537 GCTCAAAAAGAGGCTGCAAGGAGGAA-CCCTCTAAAACACACACACACACACACACACACAC 595
 QY 801 rProGlyLysArgAspGlnAlaAspGlySerLysHisLeuProGlyLysPheSerPhe 820
 DB 596 TATTGAAGGATCATGATGAAGCAGCAATGGAATGAATGAATGAATGAATGAATGAATGAATGAAT 655
 QY 821 rCysGlnLysLeuAlaLeuMetThrArgThrArgAlaArgMetLeuLysGlnLysMetAsn 840
 DB 656 CAGGCGCAAACTGCGGAAATGACTCTTATTCAAAAAATAATGCTAAAGTAAAGTAAAGTAAAG 715
 QY 841 AspSerMetAspThrSerAsnLysGlnLys 851
 DB 716 GGAAGCTGGCTAGCTTCCACACACAGGCAAAAGAG 748

Search completed: January 19, 2003, 05:05:46
 Job time : 1895.76 secs

Genware version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

EM protein multiple search, using frame_plus.p2n model

Run on: January 17, 2003, 18:27:24 ; Search time 2054.94 Seconds
(without alignments)
7420.587 Million cell updates/sec

Filter: us-09-026-459a-2
Perfect scores: 4/47
Sequences: 1 MPKRTPTTAAIAAAAAAP.....TPM_KLPMNMISMTNKEEK 928

Scoring table:
BLOSUM62
Ynapop 10.0 ; Ynapext 0.5
Ynapop 10.0 ; Ynapext 0.5
Ynapop 6.0 ; Ynapext 7.0
Ynapop 6.0 ; Ynapext 7.0

Scorched: 14154666 seqs, 8047744376 residues

Total number of hits satisfying chosen parameters: 42408142

Minimum db seq length: 0

Maximum db seq length: 2090000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 4% summaries

Command line parameters:

Model: frame_plus.p2n model -DBW-xlib
Q: us-09-026-459a-2, seqs, 8047744376 residues, 14154666 seqs, 8047744376 residues
DB-EST: QEMT-tastap -SUFFIX=1 -MINMATCH=0 -1 -LOCAL=0 -DEEXT=0
UNITS-hits START=1 END=1 MATRX=blosum62 TRANS=blosum62 -LIST=45
LOCAL=200 -THP_SCORE=pt -THP_MAX=100 -THP_MIN=0 -ALLEN=15 -MODE=LOCAL
-CUTOFF=pt -NORM=ext -HEAPSIZE=500 -MINFN=0 -MAXFN=200000000
USP-US09026459 -CGN-1.1 -HITZ=400000000 -12148 -24322 -NFP 6 -FPG 3
NO XLPXY NO MMAP LARGJQUERY -NBO_SCORES=0 -WAIT -ONCLOG -DEV_TIMESTAMP=120
WARN TIMEOUT 0 -THREAS=1 -XAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
YAPOP=10 YAPEXT=0.5 -DEPOP=6 -DEEXT=7

Database :

EST:
1: cm estbar:
2: cm estbar:
3: cm estbar:
4: cm estbar:
5: cm estbar:
6: cm estbar:
7: cm estbar:
8: cm estbar:
9: cm estbar:
10: cm estbar:
11: cm estbar:
12: cm estbar:
13: cm estbar:
14: cm estbar:
15: cm estbar:
16: cm estbar:
17: cm estbar:
18: cm estbar:
19: cm estbar:
20: cm estbar:
21: cm estbar:
22: cm estbar:
23: cm estbar:
24: cm estbar:
25: cm estbar:
26: cm estbar:
27: cm estbar:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	1894	49.5	1364	11	AK011246
2	1551	32.3	1051	13	BM561183
3	1448	30.2	881	14	BC220275
4	1349	27.9	1001	13	BM475603
5	1254	26.1	727	9	AL597811
6	1246.5	25.8	889	14	BC220275
7	1208	25.2	845	13	BM454244
8	1111	24.2	1024	13	BM450041
9	1101	24.0	842	9	AL66048
10	1056.5	22.0	914	9	AA763411
11	1043	21.7	792	9	AL564759
12	1037	21.6	695	9	AL599633
13	1030	21.5	694	10	BE982846
14	1016	21.2	623	10	AW584181
15	1013	21.1	821	9	AA754485
16	1004	20.9	922	12	BE257447
17	978	20.4	588	10	AW682344
18	974.5	20.3	709	9	AL646097
19	962	20.1	736	13	HL151568
20	959	20.0	594	10	AV715533
21	955	19.9	584	12	BE149050
22	947.5	19.8	734	12	BE144857
23	935	19.5	539	9	AL692790
24	913	19.0	615	9	AA959846
25	908	18.9	797	10	BE539278
26	892	18.8	1153	13	BM564327
27	882	18.4	572	9	AL125941
28	879.5	18.3	734	13	BL1526982
29	869.5	18.1	625	12	BE244388
30	858	17.9	597	12	BE931236
31	844	17.6	484	10	BE168095
32	833.5	17.4	646	13	BM490416
33	824	17.2	494	10	BE502897
34	823	17.2	545	10	BE692229
35	815.5	17.0	522	9	AA972785
36	802	16.7	643	13	BL1553947
37	784	16.3	2084	13	BM456728
38	751	15.7	454	10	AW503514
39	714	14.9	529	9	AA959709
40	703.5	14.7	472	9	AA955457
41	700	14.6	462	9	AL606529
42	691	14.4	514	12	BE418684
43	683	14.2	438	12	BE827847
44	681.5	14.2	799	14	CR2304
45	667	13.9	651	13	BL169641

ALIGNMENTS

RESULT 1
AK011246
LOCUS
DEFINITION
AK011246
MUS musculus 10 days embryo whole body cDNA, RIKEN full length enriched library, clone:260617M22:ref1001astoma 1, full insert sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
AK011246.1 GI:12847243
HTC: CAP trapper.
MUS musculus (strain:157BL/6J) 10 days embryo cDNA to mRNA.
clone:110:RIKEN full-length enriched mouse cDNA library
ORGANISM
MUS musculus
Eukaryota; Chordata; Mammalia; Eutheria; Rodentia; Muridae; Mus.

DB 123 CTGGACGCTGGCTGGTATGACGAGTAAAGCAAGCAATGCTGTAATGCAAGCAAGCAAGCA 782
 QY 468 LeuSerThrGluAspGluSerThrLysLeuLeuAspAsnThrPheHisMetSerLeuL 487
 DB 285 TATGCAATCAAAATTTTGAAGCAATTTGGAATGACGACATTTTTCATATGCTTTAT 842
 QY 489 GAAAGCAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 499
 DB 845 TGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 880
 RESULT 4
 BWA75603 1001 bp mRNA Linear EST 05-FEB-2002
 DEFINITION AGEN000161_640000? NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575683
 5' mRNA sequence.
 ACCESSION BWA75603
 VERSION BWA75603.1 GI:18524645
 KEYWORDS EST
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1001)
 NIH MGC <http://www.ncbi.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 UNPUBLISHED (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: rcstra@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Placed in AM1245 row 4 column 04
 High quality sequence starts 17
 High quality sequence stops 744
 Location/Qualifiers
 source
 Zorganism="Homo sapiens"
 ZDB_XREF="ZANAL:2606"
 Zclone="IMAGE:5575683"
 Zclone_lib="NIH_MGC_92"
 Zissue_type="embryonal carcinoma, cell line"
 Znote="btl6B (phase-resistant)"
 Znote="Genetic testis: Vector: pCMV sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; cloned in primed.
 Average insert size 2.5 kb. library enriched for
 full length clones and constructed by life technologies.
 Note: this is a NIH_MGC library."
 BASE COUNT 449 a 176 c 186 g 288 t
 ORIGIN
 Alignment Scores:
 Seq. No. 4 146 124 Length: 1001
 Scores 1439.00 Matches: 281
 Percent Similarity: 94.67% Conservatism: 10
 Best Local Similarity: 90.65% Mismatch: 10
 Query Match: 27.91% Indels: 9
 Gaps: 3
 us-09-026-459a-2 (1-928) x BWA75603 (1-1001)
 QY 135 LeuLysGluThrLeuSerThrLysValAspAsnAlaMetSerArgIleLeuLysLys 154
 DB 48 GTCGACGTAATTCATACATGATCAATGAAAGTGAATATGCTGCAAGCAAGCAAGCAAG 97
 QY 155 TCAAspValThrGluThrGluThrGluThrGluThrGluThrGluThrGluThrGluThr 174
 DB 98 TATGATGATGTTGTCATCTTACGAAATTCGAAAGCAAGCAAGCAAGCAAGCAAGCAAG 157
 QY 174 GthrGlnProSerSerSerThrThrGlnThrGlnThrGlnThrGlnThrGlnThrGlnThr 194

DB 158 GACACAAACCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 217
 QY 194 rTrrPleThrPheLeuLeuAlaLysGlyGluValLeuGluMetGluAspLeuValL 214
 DB 218 TTGGATTCACATTTTATAGCTAAAGGCAAGCAATGCAAAAGCAAGCAAGCAAGCAAG 277
 QY 214 cSerPheGlnLeuMetLeuGlyValLeuAspThrPheThrGlnThrGlnThrGlnThr 244
 DB 278 TTCAITTCAGTTAATGCTATGCTGCTTGAATATTATTATTAAATCTCTCTCTCACT 447
 QY 234 cLeuLysGlnThrLysThrAlaValThrProThrAsnGlySerThrProThrProThr 254
 DB 338 GGTCAAAAGCAATATAAAGACGCTGTATACCTGATTAAGCTGATCTGCAATGCAATGCA 497
 QY 254 qArqGlyGlnAspArqSerAlaArqIleAlaLysGlnLeuGluAspThrAlaGlnL 274
 DB 398 GCGACGTCAGAACAGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 457
 QY 274 cGluValLeuGlySerGlnHisGlnGlySerGlnHisGlnGlySerGlnHisGlnGly 294
 DB 458 TGAATTTCTCTGTAAGAAACATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 517
 QY 294 SASpThrThrProPheMetAsnSerLeuGluValThrSerAsnGlyLeuProGluVal 314
 DB 518 AAAATTTTATACGCTTTTATCAATTCCTTGGACCTTGTAAACATCTAATGCTATGTA 577
 QY 314 LgluAsnLeuSerLysArqTyrrGluGlnIleTyrrLeuLysAsnLysAspLeuAspAla 334
 DB 578 TCAAAATCTTCTAAACATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 637
 QY 334 qGluThrLeuAspHisAspGlyThrLeuGlnThrAspSerThrGlnThrGlnThr 354
 DB 638 ATTATTTTGTAT 697
 QY 354 rArqThrProArqLysSerAsnLeuAspThrValAspValThrProThrHisThrPy 374
 DB 698 GAGACACACCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 757
 QY 374 oValArqThrValMetAsnThrThrGlnThrLeuMetThrIleLeuAsnSerAlaSerAs 394
 DB 758 AGTAGGACGCTGATGCAATACATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 817
 QY 394 pGlnProSerGluAsnLeuThrSerThrPheAsnGlySerThrValAsnProLysGln 413
 DB 818 TCAAGCTTCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 877
 QY 413 userIleLeuLysArq--ValLysAspThrGlyThrThrPheLysGlnLys 429
 DB 878 AAAGCTATACGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 937
 QY 430 --PheAlaLysAlaValGly 435
 DB 938 TTTGCTGCTAAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
 RESULT 5
 AL597811 727 bp mRNA Linear EST 14 AUG 2001
 DEFINITION DKFZP4340816.1_1 43 (576-676) K13022 Homo sapiens cDNA clone
 DKFZP4340816.5, mRNA sequence.
 ACCESSION AL597811
 VERSION AL597811.1 GI:15160502
 KEYWORDS EST
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 727)
 AUTHORS Rouska, A., Wellenreuther, R., Mewes, H.W., Well, B., and Wiemann, S.
 TITLE EST (Rouska, A., Wellenreuther, R., Mewes, H.W., Well, B., and Wiemann, S.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Rouska, A.J.

Department Lehrach
Max-Planck-Institute for Molecular Genetics
Ihne-Strasse 73, 14195 Berlin, Germany
Tel: +49-30-84131623
Fax: +49-30-84131128

Email: pousitka@mpg-berlin.dahlem.mpg.de
This is the 5' sequence of the clone insert
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wilmann@dkfz-berlin.de
sequenced by DKFZ (German Cancer Research Center,
Heidelberg-Germany) within the cDNA sequencing consortium of the
German Genome Project.

No sl sequence available.
This clone (pK2p3130916) is available at the RZPD in Berlin.
Please contact the RZPD, Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers

1..727

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="pK2p3130916"

/clone_lib="313 (Synonym: h102)"

/cov_stage="adult"

/lab_host="R100"

/note="Vector: pRipEx2; Site_1: SP1A; Site_2: SP1B;

cDNA-collection"

BASE : UNI 248 a 161 c 130 g 198 t

ORIGIN

Alignment Scores:

Pred. No.: 8,19e-115 Length: 727
Score: 1254.00 Matches: 241
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 1
Query Match: 26.14% Indels: 0
Gaps: 0

US 09-026-459a 2 (1-928) x AL597811 (1-727)

Q7 433 TSSSThrSerLeuSerLeuPheThrLysValArgLeuAlaThrLeuArgLeu 662
|||||
14 2 AAATGATCTCTTCACGCTTTATATAAAGTATGCGTATGCTATCTGCTA 61
Q7 664 AsnThrLeuCysGluArgLeuSerGluHisProGluLeuGluHisLeuThr 682
|||||
16 6 AATAACATCTGCAACGCTCTCTCTGACACCCCAATATAGAACATATCTGAC 121
Q7 685 LeuPheGluHisThrLeuGluAspGluThrGluLeuMetArgAspArgLeuAspGlu 702
|||||
18 12 CTTCCTGACACACCTGACAAAGATATGATGATATGACAAATATGACCA 181
Q7 703 MetMetCysSerMetThrGluThrCysLysValLysAsnLysAspCulysThrLys 722
|||||
18 12 ATATGATGCTCTGATGATGCTATATGATGATGATGATGATGATGATGATGAT 241
Q7 724 IleIleValThrAlaThrLysAspLeuProHisAlaValGluGluThrPheLysArgVal 742
|||||
18 2 ATCATCTGAACAGACATCAACAGATCTCTCTGATGATGATGATGATGATGAT 301
Q7 743 LeuIleLysGluGluGluThrLysSerIleLeuValPheThrAsnSerValPheMetGlu 762
|||||
18 302 TTGATCAAAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 361
Q7 764 ArgLeuLysThrAsnIleLeuGluThrAlaSerThrArgProThrLeuSerProThr 782
|||||
18 46 ATATGAAAGAAATATTTTCTATGATGATGATGATGATGATGATGATGATGAT 421
Q7 783 ProHisIleProArgSerProThrLysPheProSerSerProLeuArgLeuProGly 802
|||||
18 422 CCGACATCTGCAACGCTCTCTCTGACACCCCAATATAGAACATATCTGAC 481
Q7 803 AsnThrLysThrSerProThrLysSerProThrLysSerProThrLysSerProThr 820
|||||

Db 482 AACATCTATATTTTCAGCCCTGAGAGAGTCCATATAATAATTTTCAGAGAGTCTGTCAGACATCA 541
Q7 823 ThrLysMetThrThrGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 842
Db 542 AAAAAATGAT 601
Q7 843 GluLysPheThrLysIleAspGluMetValCysAspSerAspArgValLeuLysArgSer 862
146 802 CAGAGAT 941
Q7 943 AAGCTCTGCTGAT 982
Db 662 GCTGAG 721
Q7 883 AspGlu 884
|||||
Db 722 CATCAA 727
RESULT 6
BQ222227
LOCUS BQ222227
DEFINITION 5', mRNA sequence.
ACCESSION BQ222227
VERSION BQ222227.1 GI:20409327
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 880)
AUTHORS NIH MGC <http://mgi.nlm.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstra@nsl.nih.gov
Tissue Procurement: AICC
cDNA Library Prepared by: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNI)
DNA Sequenced by: Advanced Biotechnology Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLNL3279 Row: 1 Column: 20
High quality sequence stop: 606.
Location/Qualifiers
FEATURES
SOURCE
1..880
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6041563"
/clone_lib="NIH_MGC_92"
/tissue_type="Embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organs: Testis; Vector: pCMV-SPT66; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 321 a 148 c 152 g 259 t
ORIGIN
Alignment Scores:
Pred. No.: 6,11e-113 Length: 880
Score: 1246.50 Matches: 255
Percent Similarity: 92.99% Conservative: 6
Best Local Similarity: 90.75% Mismatches: 13
Query Match: 25.78% Indels: 7
Gaps: 4
US-09-026-459a 2 (1-928) x BQ222227 (1-990)
2 157 ValLeuPheAlaLeuPheSerLysLeuGluArgThrCysGluLeuIleThrLeuThrGlu 176
|||||


```

650 AspSerIleIleValPheThrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeu 769
      |||
608 CAlIcIAIAlAGIAlCTAlAACTCCGCTCTTCATCCAGACAGIAGAAACAATATTTG 567
      |||
670 GlnIyAlaSerThrArgProPro ThrLeuSerProIleProHisIle-ProArgSe 788
      |||
668 CAGTAIGCTTCAGACAGCGCGCTAGCGTGTGTCACCAANTAGCTCACAATTCTCTGGA 727
      |||
788 rProIyLys---PheProSorSorProLeu---ArgIlePro---GlycIyAsn----- 803
      |||
728 GCGCTTAAAGAGIITTCCTTACCTTTCACGAAATCCGCGGAGAAACCATCC 787
      |||
804 -IleTyrlleSerProLeuLys 810
      |||
788 IAlAlITTCACCCCTCGGAA 809

RESULT 8
RM450031 1023 bp mRNA linear EST 05-FEB-2002
DEFINITION A5EN00PT 6303544 NIH MGC 72 Homo sapiens cDNA clone IMAGE:5528637
5' UTR mRNA sequence.
AUTHORS RM450031
VERSION RM450031.1 GI:18499071
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1023)
TITLE NIH-MGC http://mgs-b1.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: eugaps@mail.nih.gov
Tissue Procurement: ATCC/DCI/OLP
cDNA Library Preparation: Life Technologies, Inc
DNA Sequencing by Applied Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plates: LLAM2204 row 3 column 22
High quality sequence stop: 643.
FEATURES
Location/Qualifiers
1..1023
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5528637"
/clone_lib="NIH-MGC-72"
/tissue_type="melanotic melanoma"
/label_host="DH10B (phage-resistant)"
/label="Human skin, Ventral, PCMV SMOG6, Site_1, Ref.
Site_2, Salt; Cloned unidirectionally, Primary Oligo dT,
Average insert size 2 kb. Library constructed by Life
Technologies."
BASE COUNT 444 a 211 c 197 g 281 t
GC=31.2%

Alignment Scores:
Prod. No. 2 726-100 Length 100%
Score 1111.00 Matches 216
Percent Similarity 100.00% Conservation 0
Best Local Similarity 100.00% Mismatches 0
Query Match 23.16% Indels 0
Pos 13 Gaps 0

670 LysValLysAsnIleValPheThrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeu 732
      |||
668 CAGTAIGCTTCAGACAGCGCGCTAGCGTGTGTCACCAANTAGCTCACAATTCTCTGGA 727
      |||
788 IAlAlITTCACCCCTCGGAA 809

RESULT 9
RM450038 832 bp mRNA linear EST 29-APR 1999
DEFINITION IMAGE:1247610 is similar to gp.M25301 Mouse retinoblastoma
susceptibility protein (MOUSE);, mRNA sequence.
AUTHORS AT646038
VERSION AT646038.1 GI:4724513
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 832)
TITLE NIH-3T3 http://www.ncbi.nlm.nih.gov/ncicqap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: eugaps@mail.nih.gov
This clone is available publicly, free through DMC, contact the
IMAGE Consortium (image.llnl.gov) for further information.
MGI:661298
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: polyI not found
High quality sequence stop: 482.
FEATURES
Location/Qualifiers
1..832
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1247610"
/clone_lib="Soares mammary gland_NLMG"
/sex="female (lactating)"

```

```

62 CAlIcIAIAlAGIAlCTAlAACTCCGCTCTTCATCCAGACAGIAGAAACAATATTTG 141
      |||
753 IleValPheIyTrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeuGlnIyAla 772
      |||
122 ATAGTATTCATATACTGCTGCTTTCATGCGAGACTGAAACAAATATTTTGCAGTATGCT 181
      |||
773 SerThrArgProThrLeuSerProIleProHisIleProArgSerProIyLysPhe 792
      |||
182 TCCACCAAGCGCGCTAGCGTGTGTCACCAANTAGCTCACAATTCTCTGGA 243
      |||
793 ProSerSerProLeuArgIleProGlyValAsnIleIyIleSerProLeuLysSerPhe 812
      |||
242 CAlIcIAIAlAGIAlCTAlAACTCCGCTCTTCATCCAGACAGIAGAAACAATATTTG 301
      |||
813 TyrlsIleSerGluGlyLeuProThrProThrLysMetThrProArgSerArgIleLeu 832
      |||
302 IAlAlAAATTCAGAAAGCTGTGTAAACAAATAAATAATGACTCCAAATCAAGATCTTA 361
      |||
833 ValSerIleGlyGluSerPheGlyThrSerGluLysPheGluLysIleAsnGlnMetVal 852
      |||
362 GAlcAAATGCTGGAATCAITCGGAGACIICGAGAGAGTCCAGAAAAAATAACAGATGGA 421
      |||
853 CysAsnSerAspArgValIleLysAsnSerAlaGlyGlySerAspProGlyLysProLeu 872
      |||
422 TGTAATATCTATCTTATATTAAAGATCATGATGAAATGAGATGGAATCAATCTCGCA 481
      |||
873 TyrlsIleGlyGluSerPheGlyThrSerGluLysPheGluLysIleAsnGlnMetVal 892
      |||
482 AAAAAATATGCTTATATTAAAGATCATGATGAAATGAGATGGAATCAATCTCGCA 541
      |||
893 GlyGluSerLysPheGlnGlnLysLeuAlaGluMetThrSerThrArgThrArgMetGln 912
      |||
542 GAGAGTCTTAAATATCTTATATTAAAGATCATGATGAAATGAGATGGAATCAATCTCGCA 601
      |||
913 TyrlsIleLysMetAsnAspSerMetAspThrSerAspGlyGlyGluGlyLys 928
      |||
602 AAGATATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649

RESULT 9
AT646038 832 bp mRNA linear EST 29-APR 1999
DEFINITION IMAGE:1247610 is similar to gp.M25301 Mouse retinoblastoma
susceptibility protein (MOUSE);, mRNA sequence.
AUTHORS AT646038
VERSION AT646038.1 GI:4724513
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 832)
TITLE NIH-3T3 http://www.ncbi.nlm.nih.gov/ncicqap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: eugaps@mail.nih.gov
This clone is available publicly, free through DMC, contact the
IMAGE Consortium (image.llnl.gov) for further information.
MGI:661298
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: polyI not found
High quality sequence stop: 482.
FEATURES
Location/Qualifiers
1..832
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1247610"
/clone_lib="Soares mammary gland_NLMG"
/sex="female (lactating)"

```



```
QY 555 IsAcuIleMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuIleLysG 575
16 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575
264 AICCAATCAAGCAATGCTTCCATGCGCTTCCAGATACACCTTTATTTCACTCAATAAGC 205
QY 576 IsSerLysAspArgGlyGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuP 595
16 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596
244 ATTCTAAAGGATGCAAGAGACT--GATAAAGCTTGAAGCTGTGTGTCTCTCAAGCTGG 148
QY 597 ToleuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysL 615
16 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617
147 GTCCTCAGGTAAGCATACTGACACAGATATGTAATCTTCTCTCTAAGAATCTCAAGA 88
QY 618 ySLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerA 635
16 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637
87 AAAAGAACATCCACACACAGGCTAAATCTCCCTGCAAAATACAGACACACACAGCAAGCTCAG 28
QY 645 IaPheGlaThrGlnLysProLeuLys 643
16 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664
27 GATTGCAATGCTCAGAGGCAATGAAA 2
```

Search completed: January 19, 2003, 05:05:07
Run time : 2072.11 secs

GenCore version 5.1.1.3
Copyright (c) 1994 - 2003 CompuGen Ltd.

em nucleic nucleic search, using sw model

em em: January 16, 2003, 15:29:22 : Search time 2029.61 Seconds
(without alignments)
18469,506 Million cell updates/sec

Title: US 09 026 459A-28

Perfect score: 3455

Sequence: 1 GCGGCAACAGCAACACGCG.....AAATGAGGATTATTCATACG 3455

Scoring table: IDENTITY NDC

Gapop 10.0 , Gapext 1.0

Searches: 16354066 seqs, 809774 (76) residues

Total number of hits satisfying chosen parameters: 42308132

Minimum hit seq length: 0

Maximum hit seq length: 20000000

Post processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Database: EST

- 1: em_estb1*
- 2: em_estb1*
- 3: em_estb1*
- 4: em_estb1*
- 5: em_estb1*
- 6: em_estb1*
- 7: em_estb1*
- 8: em_estb1*
- 9: qb_est1*
- 10: qb_est2*
- 11: qb_est3*
- 12: qb_est4*
- 13: qb_est5*
- 14: qb_est6*
- 15: em_estb1*
- 16: em_estb1*
- 17: qb_est1*
- 18: em_estb1*
- 19: em_estb1*
- 20: em_estb1*
- 21: em_estb1*
- 22: em_estb1*
- 23: em_estb1*
- 24: em_estb1*
- 25: em_estb1*
- 26: em_estb1*
- 27: em_estb1*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	929.4	26.9	1024	AK011246	AK011246 Mus musculus
2	910.2	26.3	1024	BM450941	BM450941 AGENCGOIRT
3	909.4	26.3	1051	BM561083	BM561083 AGENCGOIRT
4	866	24.8	1081	BQ439221	BQ439221 AGENCGOIRT
5	844.8	24.5	881	HQ220275	HQ220275 AGENCGOIRT
6	786	22.7	1001	BM475603	BM475603 AGENCGOIRT

7	752.4	21.8	1122	13	HM545881
8	725.4	21.0	727	9	AL597611
9	720.4	20.9	889	14	BQ222227
10	699.6	20.2	824	12	BC610651
11	698.2	20.2	792	9	AL564757
12	692.2	20.0	767	12	BC576456
13	691.8	20.0	972	13	HM466194
14	678	19.6	835	13	HM454724
15	660	19.1	672	13	B1868359
16	643.6	18.6	797	10	BE549278
17	642.2	18.6	681	12	BC616219
18	617	17.9	618	13	HM264009
19	614	17.8	694	10	BE082846
20	605.2	17.5	624	10	AW584181
21	605	17.5	605	9	AL599633
22	596.4	17.3	922	12	BC254543
23	580	16.8	580	9	AL599105
24	575.2	16.6	914	9	AA763411
25	566.4	16.4	588	10	AW368234
26	558	16.2	871	9	AA763485
27	552	16.0	832	9	A1646048
28	530.2	15.3	598	10	AW582967
29	530	15.3	539	9	AL692790
30	528	15.3	572	9	AL125941
31	516.8	15.0	597	12	HE941236
32	513.2	14.9	594	10	AV715533
33	512.2	14.8	734	12	BE144857
34	505.2	14.6	736	14	B1151568
35	493.8	14.3	525	9	AA258255
36	486.4	14.1	584	12	BC149050
37	484.4	14.0	494	10	AW502887
38	481.4	13.9	484	10	HE168095
39	480.2	13.9	659	14	B1821116
40	476.6	13.8	2083	13	BM456728
41	474.4	13.7	734	14	B1526982
42	472.4	13.7	709	9	A1646097
43	459.2	13.3	497	12	BE934784
44	457	13.2	457	9	A1094215
45	457	13.2	732	9	AL598766

ALIGNMENTS

RESULT 1	AK011246	Mus musculus	10 days embryo whole body cDNA, KIKEN full length	1364 bp	mkna	linea	HTC 19 JAN 2002
LOCUS	AK011246						
DEFINITION	AK011246						
ACCESSION	AK011246						
VERSION	AK011246.1	6112847243					
KEYWORDS	HTC; CAP trapper						
SYNOPSIS	Mus musculus (strain J55676.6) 10 days embryo whole body cDNA, clone LIBRIKEN full length enriched mouse cDNA library						
ORGANISM	Mus musculus						
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus						
AUTHORS	Carninci, P., and Hayashizaki, Y.						
JOURNAL	High-efficiency full length cDNA cloning						
MEDLINE	99279253						
PUBMED	10349636						
REFERENCE 2							
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itabashi, M., Kato, H., Ozaki, Y., Morimatsu, M. and Hayashizaki, Y.						
TITLE	Normalization and subtraction of cap-trapper selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes						
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)						
MEDLINE	20493374						
PUBMED	11042159						


```

Db 861 GTTATTTATACCAATCAAAATCTCTGCTTAAATCTGCTCAATTAATAAAATCTGTACACAA 900
QY 2941 TTTGTTTCTCTTCACAAATTAATTCCTGCTGCTTTATCTGATAGTAGATGATGCTTCT 2996
Db 901 TTTGTTTCTCTTCACAAATTAATTCCTGCTGCTTTATCTGATAGTAGATGATGCTTCT 960
QY 2997 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3056
Db 961 AAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 3057 A 3057
Db 1021 A 1021

RESULT 4
LOCUS BM561083
DEFINITION BM561083 1051 bp mRNA linear EST 20-FEB-2002
5', mRNA sequence.
VERSION BM561083
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1051)
AUTHORS NIH-MGC http://mml.nhlbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cqbbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2263 row: m column: 18
High quality sequence stop: 737.
LOCATION/Qualifiers
1: 1051
SOURCE
Organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5550905"
/clone.lib="NIH-MGC-67"
/tissue.type="retroblastoma"
/lab_host="DH10B (phage-resistant)"
/notes="green eye; Vector: pCMV-Sp6T6; Site: 1 N.T.L.
Site_2: Salt; cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
BASE COUNT 221 c 197 g 299 t 4 others
ORIGIN
Query Match 26.38; Score 909.4; DB 13; Length 1051;
Best Local Similarity 98.44; Pred No 1 6e-166;
Matches 948; Conservative 6; Mismatches 13; Indels 2; Gaps 2;

QY 1261 AAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1 ATATCTTGGGATAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 40
QY 1261 TTTGATACCGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 61 TTTGATACCGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 1421 AATTTTACGAAAATTTCTGATATGAAAATTTTCTGATATGCTGCTGCTGCTGCTGCTGCT 1380
Db 121 AATTTTACGAAAATTTCTGATATGAAAATTTTCTGATATGCTGCTGCTGCTGCTGCTGCT 180

```

```

QY 1381 GAGTCTGATGGCCATATATATGTAAGAGACATCTCAGAAATCTTCATCTGACACACAT 1440
Db 181 GAGTCTGATGGCCATATATATGTAAGAGACATCTCAGAAATCTTCATCTGACACACAT 240
QY 1441 TTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 241 TTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 1501 ATCTGAATTTTATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 301 ATCTGAATTTTATCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 1561 GGAATGCAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1620
Db 361 GGAATGCAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 420
QY 1621 GTTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
Db 421 GTTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 1681 GTTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1740
Db 481 GTTATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540
QY 1741 TCTCCAAAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 541 TCTCCAAAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 1801 GAAATGCAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1860
Db 601 GAAATGCAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 660
QY 1861 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1920
Db 661 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 1921 GATATGCAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1980
Db 721 GATATGCAATATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 780
QY 1981 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
Db 781 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 840
QY 2041 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2099
Db 841 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 900
QY 2100 TCTCCAAAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2151
Db 901 TCTCCAAAGAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954

RESULT 4
LOCUS BM439221
DEFINITION BM439221 1081 bp mRNA linear EST 24-MAY-2002
5', mRNA sequence.
ACCESSION BM439221
VERSION BM439221.1 GI:21178297
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1081)
AUTHORS NIH-MGC http://mml.nhlbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cqbbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.

```



```

107 TAGAAATGATACAGAAATTAATCTCTCTTAAGAACATGATGATTAATATAGATG 766
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: IMAGE12326 row: f column: 14
High quality sequence start: 17
High quality sequence stop: 734.
Location/Qualifiers
Source
1. .1001
/organism "Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIM_MGC_92"
/lab_host="DH10b (phage-resistant)"
/issue_type "embryonal carcinoma, cell line"
/Note "Organ: Testis; Vector: pCMV-Sport6; Site: L. Noli;
Site_2: Sali; Cloned unidirectionally; oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by life technologies.
Note: this is a NIM_MGC library."
BASE COUNT 349 a 178 c 186 g 288 t
ORIGIN
Query Match 22.7%; Score 786, DB 13, Length 1001;
Best local Similarity 95.1%; Pred. No. 1, Id: 112;
Matches 833; Conservative 0; Mismatches 40; Indels 3; Gaps 2;
QY 315 AGAAATTCATACAGTACCAATTCATATGCTATGTCAGACTGTCGAAAGAAATATGA 174
DB 43 AGTAATTCATACAGTACCAATTCATATGCTATGTCAGACTGTCGAAAGAAATATGA 102
QY 375 TGTATTTTTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 433
DB 103 TGTATTTTTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 162
QY 434 AACCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 493
DB 164 AACCCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 222
QY 444 TGCATTTTTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 553
DB 223 TGCATTTTTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 282
QY 554 TGCATTTTTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 613
DB 283 TGCATTTTTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 342
QY 614 AAGAACCATATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673
DB 343 AAGAACCATATAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 402
QY 674 GTCACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 733
DB 403 GTCACAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 462
QY 734 TTTCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 793
DB 463 TTTCTCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
QY 794 TTATACCTTTTATGAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 853
DB 523 TTATACCTTTTATGAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 582
QY 854 ATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 913
DB 583 ATCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 642
QY 914 TTTTGATATATGAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 973
DB 643 TTTTGATATATGAAATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 702
QY 974 CATTACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1033
DB 703 CATTACCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 762

```

```

FEATURES
Source
1. .1001
/organism "Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIM_MGC_92"
/lab_host="DH10b (phage-resistant)"
/issue_type "embryonal carcinoma, cell line"
/Note "Organ: Testis; Vector: pCMV-Sport6; Site: L. Noli;
Site_2: Sali; Cloned unidirectionally; oligo-dt primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by life technologies.
Note: this is a NIM_MGC library."

```

```

RESULTS
BM475603
DEFINITION
BM475603
VERSION
BM475603.1 GI:18524645
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1001)
RefSeq http://www.ncbi.nlm.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rstraus@mail.nih.gov
Tissue Procurement: AACC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be

```

QY 1044 GGAATGTTATGAAACATATGAAACATATGATGATTTTAAATTCAGAGAAATGATCAAA 1094
 DB 764 GCAATGTTATGAAACATATGAAACATATGATGATTTTAAATTCAGAGAAATGATCAAA 822
 QY 1094 GGAATGTTATGAAACATATGAAACATATGATGATTTTAAATTCAGAGAAATGATCAAA 1151
 DB 824 GGAATGTTATGAAACATATGAAACATATGATGATTTTAAATTCAGAGAAATGATCAAA 882
 QY 1192 ACTTAAAG 1187
 DB 884 GGAATGTTATGAAACATATGAAACATATGATGATTTTAAATTCAGAGAAATGATCAAA 918
 RESULT 7
 RMS 5.981 1122 bp mRNA linear EST 20-FEB-2002
 DEFINITION A26NC0001 6505143 NIH_MGC_125 Homo sapiens cDNA clone IMAGE:5588245
 5' mRNA sequence.
 ACCESSION U045981
 VERSION U045981.1 GI:18778465
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 1122)
 NIH MGC <http://mgi.nlm.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT Robert Strausberg, Ph.D.
 EMAIL: compuser@mail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)
 RNA Sequencing by: Agencourt Bioscience Corporation
 clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LAM2359 row: a column: 14
 High quality sequence start: 48
 High quality sequence stop: 773.
 Location/Qualifiers
 1..1122
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /feature="IMAGE:5588245"
 /feature="NIH_MGC_125"
 /feature="EST"
 /feature="cDNA"
 /feature="mRNA"
 /feature="linear"
 /feature="EST 20-FEB-2002"
 /feature="5' mRNA sequence"
 /feature="cDNA clone"
 /feature="IMAGE:5588245"
 /feature="NIH_MGC_125"
 /feature="EST"
 /feature="human"
 /feature="Homo sapiens"
 /feature="Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo."
 /feature="1 (bases 1 to 1122)"
 /feature="NIH MGC http://mgi.nlm.nih.gov/"
 /feature="National Institutes of Health, Mammalian Gene Collection (MGC)"
 /feature="Unpublished (1999)"
 /feature="CONTACT Robert Strausberg, Ph.D."
 /feature="EMAIL: compuser@mail.nih.gov"
 /feature="Tissue Procurement: Invitrogen"
 /feature="cDNA Library Preparation: Life Technologies, Inc."
 /feature="cDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)"
 /feature="RNA Sequencing by: Agencourt Bioscience Corporation"
 /feature="clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov"
 /feature="Plate: LAM2359 row: a column: 14"
 /feature="High quality sequence start: 48"
 /feature="High quality sequence stop: 773."

DB 460 ATTGAAGATCAGATGAAAGATGAAAGATGAAAGATGAAAGATGAAAGATGAAAGATGAAAG 419
 QY 2602 CAGAAACAGCAG 2661
 DB 420 CAGAAACAGCAG 479
 QY 2602 CAGAAACAGCAG 2721
 DB 480 AGCAAGCAGCAG 549
 QY 2722 CAGAAACAGCAG 2781
 DB 540 CAGAAACAGCAG 599
 QY 2782 CAGAAACAGCAG 2841
 DB 600 CAGAAACAGCAG 659
 QY 2842 CAGAAACAGCAG 2901
 DB 660 CAGAAACAGCAG 719
 QY 2902 CAGAAACAGCAG 2961
 DB 720 CAGAAACAGCAG 779
 QY 2962 CAGAAACAGCAG 3021
 DB 780 CAGAAACAGCAG 839
 QY 3022 CAGAAACAGCAG 3081
 DB 840 CAGAAACAGCAG 898
 QY 3082 CAGAAACAGCAG 3148
 DB 899 CAGAAACAGCAG 958
 QY 3139 CAGAAACAGCAG 3198
 DB 959 CAGAAACAGCAG 1017
 QY 3199 CAGAAACAGCAG 3257
 DB 1018 CAGAAACAGCAG 1077
 QY 3258 CAGAAACAGCAG 3317
 DB 1078 CAGAAACAGCAG 1137
 RESULT 8
 AL597811 727 bp mRNA linear EST 14 AUG 2001
 DEFINITION DKE261400816_21 213 (synonym: hicc-2) Homo sapiens cDNA clone
 LECHUS
 ACCESSION AL597811
 VERSION AL597811.1 GI:15160502
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 1 (bases 1 to 727)
 Poustka, A., Wellenreuther, R., Mewes, H.W., Wolf, R., and Wiemann, S. (1999)
 EST (Poustka, A., Wellenreuther, R., Mewes, H.W., Wolf, R., and Wiemann, S.)
 Unpublished (1999)
 CONTACT Poustka, A.J.
 DEPARTMENT Lehrstuhl
 MAX-Planck-Institut für Molekulare Genetik
 Ilmenstrasse 73, 14195 Berlin, Germany

full-length clones and constructed by Life Technologies.
Note: this is a NIH/MGC Library."

```
BASE COUNT      228 a      124 c      134 g      186 t
ORIGIN

Query Match.          19.1%   Score 660;   DR 14;   Length 672;
Best Local Similarity 99.9%   Pred. No. 4.4e-118;
Matches 6/1;   Conservative 0;   Mismatches 0;   Indels 1;   Gaps 1;

QY 2321 AATATTCAGCTGCAAGAGTCATATAAAATTCACAGAGTCGCAACACCAACAAAAA 2380
DB 1 AATATTCAGCTGCAAGAGTCATATAAAATTCACAGAGTCGCAACACCAACAAAAA 60

QY 2481 GATATTCAGAGTAAATCTTACATCAATTCGTGAATCATTCGGGATTTCTGAGAAAT 2440
DB 61 GATATTCAGAGTAAATCTTACATCAATTCGTGAATCATTCGGGATTTCTGAGAAAT 120

QY 2443 TCAAGAAAAATCAATCAATGATGATGATACAGAGTCTGTGCTGCTCAAAAAGATGCTGAAG 2500
DB 121 TCAAGAAAAATCAATCAATGATGATGATACAGAGTCTGTGCTGCTCAAAAAGATGCTGAAG 180

QY 2501 CAAGCAAAAGCTGCTAAACATCTCAAAAAACATATCTTCATATTCGAAGGATCAGATCAAG 2560
DB 181 CAAGCAAAAGCTGCTAAACATCTCAAAAAACATATCTTCATATTCGAAGGATCAGATCAAG 240

QY 2561 CATATGGAAGTAAATATCTGTCAGAGAGAGTGCATAATTCAGCAGAAACTGGCAGAAATGA 2620
DB 241 CATATGGAAGTAAATATCTGTCAGAGAGAGTGCATAATTCAGCAGAAACTGGCAGAAATGA 300

QY 2621 CTCTTACTGCAAGCAGATGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2680
DB 301 CTCTTACTGCAAGCAGATGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360

QY 2681 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2740
DB 361 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

QY 2741 GTCCTTCAGATGATGATATATATCTTTCAGAGTCTCTTTATGAGCAATTTAATAT 2800
DB 421 GTCCTTCAGATGATGATATATATCTTTCAGAGTCTCTTTATGAGCAATTTAATAT 480

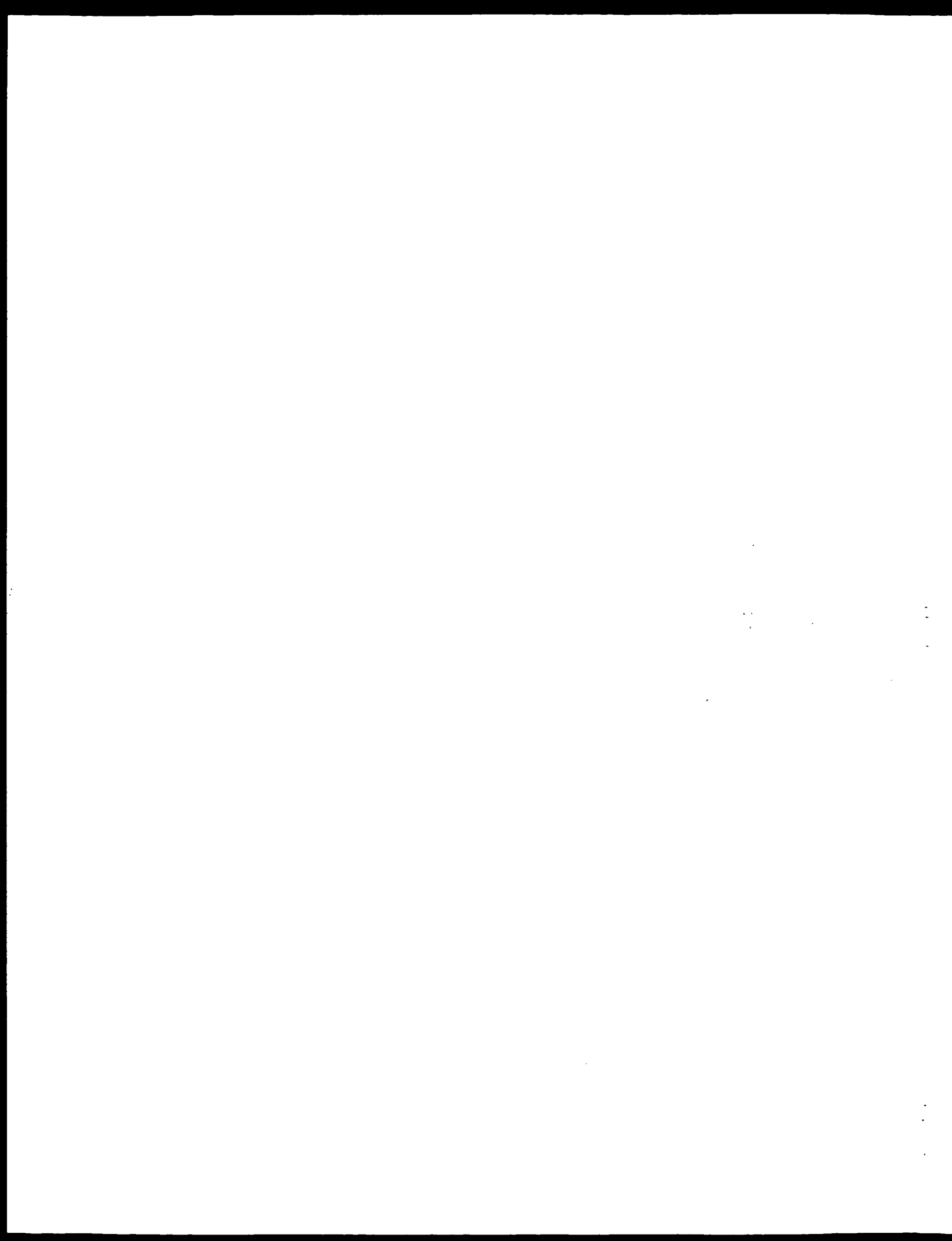
QY 2801 CTTCAGCTCTTTTGTGATATAAAATGTCAGATGCAATTTGTTGCTGATTCCTAAGC 2860
DB 481 CTTCAGCTCTTTTGTGATATAAAATGTCAGATGCAATTTGTTGCTGATTCCTAAGC 540

QY 2861 CATCTGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2919
DB 541 CATCTGAAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

QY 2920 CATTTAAAAATGTTATGAGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2979
DB 601 CATTTAAAAATGTTATGAGCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

QY 2980 ATAGTAAAGAAATG 2991
DB 661 ATAGTAAAGAAATG 672
```

Search completed: January 14, 2003, 04:06:42
Job Time: 4063.69 sec




```

121 ACATATTTTATAGTAAAGGGAAGATATACAAAGCAGACATGATGTTGATTTTCATTT 180
QY
184 GlnLeuMetLeuGysValLeuAspTyrPheIleLysLeuSerProMetLeuLeuLys 203
DB
181 CAGTTAATGATATGCTGCTGGATATTTTATTAAAGTCTCACTGGALGTTGCTCAA 240
QY
204 GlnProLysThrAlaValIleProIleAsnGlySerProArgPheProGATGAGGly 223
DB
241 GAAATATATAAAGAGTGTATATAGTATGATTAATGTTTCACTGGAAGAGAGAGGT 300
QY
224 AlaGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 243
DB
301 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY
244 LeuGysValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 263
DB
301 CTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY
264 LeuProPheMetAsnSerLeuGlyGluValIleSerAsnGlyGluProLeuVal 283
DB
421 ATAGCTTTTATGAAATCTCTGGAGCTGTAAACATCTAAAGAGCTTCAGAGGTGA 480
QY
284 LeuSerLysArgTyrGluGluIleTyrLeuLysAsnLysAspLeuAspAlaArgLeuPhe 303
DB
481 CTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY
304 LeuAspHisAspLysThrLeuLeuThrAspSerIleAspSerPheGluIleArgThr 323
DB
541 TTGATCATGATAAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY
324 ProArgLysSerAsnLeuAspGluValIleValIleProIleProIleProIlePro 343
DB
601 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY
344 ThrValMetAsnThrIleGluGluLeuMetMetIleLeuAsnSerAlaSerAspGlu 363
DB
661 ACTGTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY
364 GserGluAsnLeuLeuSerTyrPheAsnAsnCysThrValAsnProLysGluSerI 382
DB
721 TTTAAAAATGATTTTCTTTTATTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY
382 LeuLeu-----LysArgValLysAspIleGlyTyrIlePhe-----LysGluLysPhe 397
DB
781 ATATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
RESULT 7
HM454724
Locus
DEFINITION HM454724 835 bp mRNA linear EST 05-FEB-2002
5' mRNA sequence.
ACCESSION HM454724.1 GI:18502704
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mrc.nhlbi.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rstraus@nih.gov
Tissue Procurement: ATCC/DCTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plates: 12x12x206 row: c column: 07
high quality sequence stop: 647.

```

FEATURES

```

source
location/Qualifiers
1..845
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGP:5528014"
/clone_11b="NIH-MGC_72"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage resistant)"
/notes="organ: skin; Cloned unidirectionally. Primers: cDNA dfr.
Site_2: SalI; Cloned unidirectionally. Primers: cDNA dfr.
Average insert size: 2 kb. Library constructed by Life
technologies."
BASE COUNT 254 a 147 g 244 t
ORIGIN
Alignment Scores:
Pred. No.: 2,08e-119 Length: 845
Score: 1208.00 Matches: 244
Percent Similarity: 92.91% Conserved: 5
Best Local Similarity: 91.04% Mismatches: 12
Query Match: 26.18% Indels: 7
DB: 13 Gaps: 5
US-09-026-459A-29 (1-895) x HM453724 (1-835)
QY 517 LeuGluArgCysGluHisArgIleMetGluSerLeuAlaIleProLysSerAspSerProLeu 536
DB 8 GTGGAGGATGTGAACATCGAATCATGAATCGCTTCATGGTCTCAGATTCAATTTTA 67
QY 537 PheAspLeuIleCysGluSerLysAspArgGlyProIleAspHisLeuGluSerAla 556
DB 68 TTGATCTTATTAACCAATCAAGGACGACAGACGACAACTGATCACTCAATATGAT 127
QY 557 CysProLeuAsnLeuProLeuGluAsuAsnIleThrAlaAlaAspMetTyrLeuSerPro 576
DB 128 TGTCTCTTAATCTCTCTCCAGCAATTAATCACTGACAGACATAGATCTTCTCT 187
QY 577 ValArgSerProLysLysLysGlySerThrThrArgValAsnSerThrAlaAsnAlaGlu 596
DB 188 GTAAATGCTCTCAAGAAAAAGGTTCAATACGCTGAGGTTCAATTTGATTTGAT 247
QY 597 ThrGlnAlaThrSerAlaPheGluThrGluLysProLeuLysSerThrSerLeuSerLeu 616
DB 248 ACACAGCAACCTCAGCTCCACACGACGACAAAGCCATTCAGATCTCTCTCTCT 307
QY 617 PheTyrLysLysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuGysGluArgLeu 636
DB 308 TTTTATAAAAAAGTGTATGCGCTAGCCTATCTCGGCTAAATACATCTGTTGAACT 367
QY 637 LeuSerGluHisProGluLeuGluHisIleIleIlePheLeuPheGluHisThrLeuGlu 656
DB 368 CTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 427
QY 657 AsnGlyTyrGluLeuMetArgAspArgGlySerLeuAspGlyMetMetLysSerMetTyr 676
DB 428 AATGAGTATGAGCTATCAGACAGACAGAGATTCAGCAAAATATGAGAGAGAGAG 487
QY 677 GlyIleCysLysValLysAsnIleAspLeuLysPheLysIleIleValThrAlaTyrLys 696
DB 488 GGCATATGCAAGGACAGCAATATAGAGCTTAAATTTAAAAATATGTAACAGATAC 547
QY 697 AspLeuProHisAlaValGluGluThrPheLysArgValLeuLeuLysLeuLeuLeu 716
DB 548 GATCTCTCTGAGCTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 607
QY 717 AspSerIleIleValPheTyrAsnSerValPheMetGluArgLysThrAsnIleLeu 736
DB 608 GATCTCTATATAGTATCTATTAATCTGCTGCTTCAAGAGAGAGAGAGAGAGAG 647
QY 737 GlnTyrAlaSerThrArgProProProProProProProProProProProPro 755
DB 668 CAGTATGCTTTCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 727

```

```

QY 765 pProtyrLys---PheProSerProLeu---ArgilePro---GlycIAsn----- 770
||||| 333 ||||||||| ||| ||||||| |||||||
LB 728 GAGCTTAAGATTTCCTTACAGCTTTAGAGAAATCCGAGGAAAGGAAAGATCC 787
QY 771 -ileTyrrileSerProLeuLys 777
||||| 333 ||||||||| |||
LB 748 TATAATTTTCAACCCCGCTGGAA 809

RESULT 8
BM450031
LOCUS       BM450031
DEFINITION  A-FNC-apt_6403544 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5528037
5' mRNA sequence.
VERSION     BM450031
KEYWORDS    EST
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE   1. (bases 1 to 1023)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo
TITLE       NIH MGC http://mgi.nih.gov/
JOURNAL     National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: eus@pshs@mail.nih.gov
Tissue Procurement: ATCC/DCTD/BTP
cDNA Library Preparation: Life Technologies, Inc
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Amersham Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/HLI at:
http://hgwat-llnl.gov
Plate: LLAM12204 row: 3 column: 22
High quality sequence stop: 643.
FEATURES             location/qualifiers
     source           1..1023
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5528037"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /vector="puc19"
                     /cdate="Apr 96"
                     /site="SalI; Cleaved unidirectionally. Primer: clone 8T.
                     Average insert size 2 kb. Library constructed by Life
                     Technologies."
BASE 1 UNL 334 a 211 c 197 g 281 t
ORIGIN
Alignment Scores:
Prod. No.:      8.24e-109      Length:      1023
Score:          1111.00      Matches:      216
Percent Similarity: 100.00%      Conservat: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    24.08%      Indels: 0
EB:             13          Gaps: 0

US 09-026-459A 29 (1-895) x BM450031 (1-1023)
QY 760 LysValLysAsnTicAspGluLysPheLysTicLeValThrAlaLysAspGluPro 699
||||| 333 ||||||||| ||| ||||||| |||||||
LB 728 AAGATGAGAAATATAGCTTAATTTAAATTTATCTTACACATACAGGATTTCT 717
QY 770 HisAlaValGlnGlnThrPheLysArgValLeuLysGlnGluLysTyrAspSerile 719
||||| 333 ||||||||| ||| ||||||| |||||||
LB 748 CATGCTGTCACGACATTCACAAAGCTTTTCATCAACACACAGCAGATTCAT 721
QY 740 IleValIleThrAsnSerValPheMetGlnArgLeuLysThrAsnLeuLeuLysTyrAla 739
||||| 333 ||||||||| ||| ||||||| |||||||
LB 728 ATAATATATATATATATATATATATATATATATATATATATATATATATAT 741
QY 740 SerThrArgProProThrLeuSerProTicProHisTicProArgSerProTyrLysPhe 759

```

```

LB 182 TCCACACGCGCCGCTACCTTCACCAATACCTTACATTCCTCCGAACCCCTACAACTT 241
QY 760 ProSerSerProLeuArgIleProGlyLysAsnIleTyrrileSerProLeuLysSerPro 779
||||| 333 ||||||||| ||| ||||||| |||||||
LB 242 CCAAGTTCACGCTTACGGAATCTCTGGAGGAGCAACATCATATTTTCAGCTTCGAGAGTCA 401
QY 780 TyrrLysIleSerGlnGlyLeuProThrProThrLysMetThrIleArgSerAlaGlnGlu 799
||||| 333 ||||||||| ||| ||||||| |||||||
LB 302 TATATAATTTTCAAGAGTCTGCGACACACCAACAAAATGACTCCAGATTCACAAATCTTATA 461
QY 909 ValSerTicGlyValLysSerPheGlyThrSerGlnLysGlnLysTicAsnGlnMetVal 819
||||| 333 ||||||||| ||| ||||||| |||||||
LB 362 GATCAATTCGTTCAATTCATTCGGGACTTCGACAGGTTCCAGAAAAATAATTCAGATGGTAA 421
QY 920 CysAsnSerAspArgValLeuLysArgSerAlaGlnGlySerAsnProProLeuProLeu 849
||||| 333 ||||||||| ||| ||||||| |||||||
LB 422 TGTAACTGCTTATCTGCTTAAAGAAATGCTGAAAGAAAGCAAAAGCTGCTTAAAGCAGCTG 481
QY 940 LysTyrLysLeuArgPheAspLeuGlySerAspGlnValAspGlySerTyrLysHisLeuPro 869
||||| 333 ||||||||| ||| ||||||| |||||||
LB 482 AAAAAATATAGTTTATATATTGAAAGATATATATGAAATATATATGAAATATATCTGCA 541
QY 960 GlyGlnSerLysPheGlnGlnGlySerAlaGlnMetThrSerThrArgThrArgMetGln 879
||||| 333 ||||||||| ||| ||||||| |||||||
LB 542 GCAGACTGCAAAATTCAGCAGCAAAATTCGACAGAAATGACTTCTACTCGAACACAGAAATGCAA 601
QY 980 LysAlaLysMetAsnAspSerMetAspThrSerAsnLysGlnGluLys 895
||||| 333 ||||||||| ||| ||||||| |||||||
LB 602 AAGCAGAAAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
PFSUIT 9
AA763411/C
LOCUS      AA763411
DEFINITION w54404.r1 Soares_mammary_gland_NLMNG Mus musculus cDNA clone
IMAGE:1247598 5' similar to gb:M15400 RETINOBLASTOMA-ASSOCIATED
PROTEIN (HUMAN); gb:M26391 Mouse retinoblastoma susceptibility
protein (MOUSE);, mRNA sequence.
ACCESSION  AA763411
VERSION     AA763411.1 GI:2813158
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
Eukarya; Euk; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerogathi; Muridae; Murinae; Mus.
1 (bases 1 to 914)
Marra, M., Hillier, L., Allen, M., Bowles, M., Diehrich, N., Dubouque, I.,
Geisel, S., Kucaba, T., Levy, M., Le, M., Martin, J., Mottis, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, H.,
Theisinger, R., Wyllie, T., Lennon, G., Soares, R., Wilson, R. and
Waterston, R.
The WashU-HIMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HIMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
this clone is available royalty free through HMI; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:661286
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 422.
location/Qualifiers
     source           1..914
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1247598"
                     /clone_lib="Soares_mammary_gland_NLMNG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
FEATURES
SOURCE

```

/Lab host="B1010"

/Note="Vector: pT7-6-pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - o1010(d1) primer. Double-stranded cDNA was ligated to EcoRI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7-6 vector. Library is normalized. Library was constructed by Honto Soares and M. Fatima Bonafide."

BASE COUNT 217 a 178 c 217 g 261 t 1 others

Alignment Scores:

Prod. No.: 5,25a-103 Length: 914

Score: 1056.50 Matches: 244

Percent Similarity: 82.14% Conservat: 19

Best Local Similarity: 75.97% Mismatches: 53

Query Match: 22.90% Indels: 6

DB: 9 Gaps: 2

US-09-026-459a-29 (1-895) x AA769411 (1-914)

QY 303 PheLeuAspHisAspLysThrLeuGlnThrAspSerThrAspSerPheGlnThrGlnArg 322

Db 914 TTTTGGATCCAGATTAATAGGCTGCTGCTTTTAAAGCGGGAG 857

QY 323 ThrProAlaLysSerAsnLeuAspGlnValAsnValIleProProHisThrProVal 342

Db 856 AGGATGCGAAATATATGCTTGAAGCGGCTAAATGCTGCTGCTGCTGCTGCTGCT 798

QY 343 ArgThrValMetAsnThrIleGlnIleLeuMetIleLeuAsnSerAlaSerAspGln 362

Db 797 TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 738

QY 363 ProSerGlnAsnLeuIleSerThrAspAsnGlyThrValAsnProLysGlnSerIle 382

Db 747 GCTGCAAGATTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679

QY 383 GlnLysAlaValLysAspIleGlyThrIleProLysGlnLysPheAlaLysAlaValGly 402

Db 678 TAAAGCGGCTTAAAGCGGCTTAAAGCGGCTTAAAGCGGCTTAAAGCGGCTTAAAG 623

QY 403 GlnLysValValIleGlySerGlnArgThrLysValAlaIleThrArg 422

Db 622 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 563

QY 423 ValMetGlnSerMetLysSerGlnGlnGlnArgLysSerIleGlnAsnProLys 442

Db 562 GTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 503

QY 443 LeuLeuAsnAspAsnIleProHisMetSerLeuLeuAlaLysAlaLeuValMet 462

Db 502 GCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 443

QY 463 AlaThrTrpSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro 482

Db 442 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 383

QY 483 TrpIleLeuAsnValLeuAsnLeuLysAlaIleAspThrLysValIleGlnSerPhe 502

Db 382 TGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 323

QY 503 ThrLysAlaLeuLysAsnLeuThrArgMetIleLysIleGlnAlaArgLysGlnHis 522

Db 322 AATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 263

QY 523 ArgIleMetGlnSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuLysGln 542

Db 262 GAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 203

QY 543 SerLysAspArgGlnGlyProThrAspHisLeuGlnSerAlaLysProLeuAsnPro 562

Db 202 TGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 146

QY 563 LeuGlnAsnAsnHisThrAlaAlaAspMetCytLeuSerProValAlaSerThrLysLys 582

Db 145 CTCACAGCTAACCAATACGACAGCAATACGATCTCTCTCTCTCTCTCTCTCTCTCT 86

QY 583 LysLysThrThrArgValAsnSerThrAlaAsnAlaSerThrAlaAlaThrAla 602

Db 85 AGAAATTGCAATATATGTAATATCTGTCGTAATATATATATATATATATATATAT 26

QY 603 PheGlnThrGlnLysProLeuLys 610

Db 25 TTCCATATATAGAAAGCATTCGAA 2

RESULT 10

AL563757 792 bp mRNA Linear EST 16 FEB 2001

LOCUS

DEFINITION AL563757 L11.NF1001.NM04 Homo sapiens cDNA clone (S010007YA05.4

DELINE mRNA sequence.

ACCESSION AL563757.1 GI:12913464

VERSION

KEYWORDS EST.

SOURCE human.

ORGANISM Homo Sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 792)

AUTHORS Li, W.H., Gruber, C., Jesse, J., and Pelayer, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope

Genoscope - Centre National de Sequençage

BP 191 91006 EVRY cedex - France

Email: seqgen@genoscope.cns.fr, Web: www.genoscope.cns.fr.

LOCATION/Qualifiers

1..792

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS000007YA05"

/cotton="Lib-111.NF1001.NM04"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="B1010"

/note="cDNA. Brain. Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI o1010(d1) primer. Five prime end enriched, double stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9600 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 401 610 8371 Email: fliang@lifestech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 246 a 148 c 117 g 255 t 26 others

ORIGIN

Alignment Scores:

Prod. No.: 1,210-101 Length: 792

Score: 1043.00 Matches: 217

Percent Similarity: 84.53% Conservat: 7

Best Local Similarity: 81.89% Mismatches: 48

Query Match: 22.61% Indels: 4

DB: 9 Gaps: 0

US-09-026-459a-29 (1-895) x AL563757 (1-792)

QY 207 LysThrAlaValIlePheProIleAsnGlySerArgThrProAlaGlyThrAlaAsnArg 226

Db 791 AAAATGAGCTKTATATAGCTATTAAGGTTACCTGAAATATATATATATATATATAT 742

QY 227 SerAlaArgThrAlaLysGlnIleGlnAlaAspThrArgThrIleGlnValLysLys 246

Db 741 AGTGCATGAGTATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 672

I and cloned into the Not I and Eco RI sites of the modified pT7A vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Ronaldo.

BASE COUNT 266 a 170 c 176 g 214 t 6 others

Alignment Scores:

Pred. No.: 4,32e-102 Length: 832
Score: 1029.00 Matches: 217
Percent Similarity: 88.64% Conservative: 17
Best Local Similarity: 82.28% Mismatches: 29
Query Match: 22.30% Indels: 6
Gaps: 1

US-09-026-459A-29 (1-895) x A1646038 (1-832)

```

47 2 GlnAspSerGlyProGluAspLeuProValArgGluGluGluThrGluGlu 21
14 47 CAGCAACAGCGCGGAGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 106
47 22 ProAspPheThrAlaLeuCysGlnLysLeuLysLysLysValArgGluArgAla 41
14 47 CCGCAAAATATGCGATATATGCGATATATGCGATATATGCGATATATGCG 166
47 42 TrpLeuThrTrpGluLysValSerSerValAspLysValLeuGlyGlyTrpLeuLys 61
14 47 TGGCTAACTTGGCAAAATATGCGATATATGCGATATATGCGATATATGCG 226
47 62 LysLysGluLeuThrGlyLysLysLysLysLysLysLysLysLysLysLys 81
14 47 AAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 286
47 82 PheThrPheThrGluLeuGlnLysLysLysLysLysLysLysLysLysLysLys 101
14 47 TACATCTTACATCTTACATCTTACATCTTACATCTTACATCTTACATCTT 345
47 102 LeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 121
14 47 TAAAGCAAAATGCGATATATGCGATATATGCGATATATGCGATATATGCG 406
47 124 TyrAspValLeuPheAlaLeuPheSerLysLysLysLysLysLysLysLys 141
14 47 TATATATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTTATGTT 466
47 142 ThrGlnProSerSerSerLysThrGluLeuAsnSerAlaLeuValLeuLysValSer 161
14 47 ACACAAACACACACACACACACACACACACACACACACACACACACACAC 526
47 162 TrpLeuPhePheLeuValLeuValLeuValLeuValLeuValLeuValLeu 181
14 47 TGGATATATATATATATATATATATATATATATATATATATATATATAT 586
47 182 SerPheGlnLeuMetLeuCysValLeuAspSerTrpPheLysLeuSerTrpMetLeu 201
14 47 TCAATTCAGCTAATTCAGCTAATTCAGCTAATTCAGCTAATTCAGCTAATTC 645
47 202 LeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 221
14 47 CACACACACACACACACACACACACACACACACACACACACACACACAC 705
47 222 ArgGlyValAsnArgSerAlaArgLysAlaLysLysLysLysLysLysLys 241
14 47 ACAAGCTACACACACACACACACACACACACACACACACACACACACACAC 764
47 242 TrpValLeuSerGlySerGlySerGlySerGlySerGlySerGlySerGly 261
14 47 GAGTTTCTCTTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 818
47 262 AsnPheLeuPro 265
14 47 AAAATATACCT 830

```

RESULT 14

AW583181

DEFINITION

hail09.yl Human Pancreatic Islets Homo sapiens cDNA clone

IMAGE:5638049 5' similar to gb:M15400 RITINORIASOMA-ASSOCIATED

PROTEIN (HUMAN): gb:M26391 Mouse retinoblastoma susceptibility

protein (MOUSE): mRNA sequence.

ACCESSION

AW583181

VERSION

AW583181.1 GI:7260125

KEYWORDS

EST.

SOURCE

human.

REFERENCE

AUTHORS

Melton, D., Meadows, A., Clifton, S., Hillier, L., Marra, M., Pape, B.,

Wyllie, J., Martin, J., Blistein, A., Schmitt, A., Theising, B., Kitter

, E., Renko, J., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole

, R., Tsakarelshvili, R., Williams, T., Jackson, Y., and Bowers, Y.

WashU-Harvard Pancreas EST Project

Unpublished (2000)

COMMENT

CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Indocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept. of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-9557

Email: dmelton@biohp.harvard.edu

Libraries were constructed by Dr. Douglas Melton

DNA sequencing by Washington University Genome Sequencing Center

For information on obtaining a clone please contact: Juliana Brown

(brownj@fas.harvard.edu)

This sequence now available from the IMAGE Consortium, for clone

orders contact: info@image.lln.gov

Seq primer: -40bp from Gibco

High quality sequence stop: 430.

Location/Qualifiers

1..623

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5638049"

/clone_lib="Human Pancreatic Islets"

/issue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="PH108"

/note="Organ: Pancreas; Vector: pSPOR11; Site: 1; Not 1;

Site: 2; Sal 1; Library constructed using Superscript

Plasmid Library kit (Life Technologies); cDNA made by

oligo dT priming; Size selected by column fractionation;

average insert size 1.08 kb. Primary library.

unamplified."

HASH COUNT 234 a 93 c 109 g 186 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 6.87e-99 Length: 623

Score: 1035.00 Matches: 201

Percent Similarity: 98.07% Conservative: 3

Best Local Similarity: 97.10% Mismatches: 3

Query Match: 22.02% Indels: 1

DB: 10 Gaps: 0

US-09-026-459A-29 (1-895) x AW583181 (1-623)

47 238 TATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257

14 238 TATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 257

DB 3 AAGAGATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 52

47 258 ValThrPheLysAsnThrThrPheMetAsn-SerLeuGlyLeuValThrSerAsn 277

14 258 ValThrPheLysAsnThrThrPheMetAsn-SerLeuGlyLeuValThrSerAsn 277

DB 63 GTTATTTCAAAAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 122

47 277 YLeuProGlnValGlnAsnLeuSerLysArgTrpGluThrLeuTrpLeuYAsn 297


```

UY 542 701CGGlaAsnAsuHisThrAlaAlaAspMetCysIleCysSerProValArgSerProLysL 582
146 147 GTCGCCAGGTAAACCATACCTAGCAGCAGATATGTATCTTCTCCCTAGATCCTCAGAGA 88
UY 582 ystLysGlySerThrArgValAsuSerThrAlaAsnAlaGluThrGlnAlaThrSera 602
148 87 AAAGAAGCTTCCACCTACACGCTAAATTTCTGCTGCAAAATACAGACACACAAAGCGCTCAG 28
UY 602 laPheGlnThrGlnLysProLeuLys 610
149 27 ATTCTATATCTACAGACCCATTCAGAA 2

```

Search completed: January 19, 2003, 05:05:21
 Job time : 1994.1 secs

Genotype version 5.1.1.4
Copyright (c) 1994 - 2003 CompuGen Ltd.

AM protein nucleic search, using frame_plus_pzn model

Run on: January 17, 2003, 18:27:24 : Search time 1727.82 Seconds
(without alignments)
7420.587 Million cell updates/sec

Title: us-09-026-459a-37

Pattern score: 4017

Sequence: 1 MSHIKKIVLVITATPSKTFPTPMKKPKNNISMTSNEEK 781

Scoring table:

Match 62
Xgapop 10.0 : Xgapext 0.5
Ygapop 10.0 : Ygapext 0.5
Zgapop 6.0 : Zgapext 7.0
Delop 6.0 : Delext 7.0

Scanned: 16154096 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 2 (0.00012)

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0K

Maximum Match 10K

Listing first 45 summaries

Command line parameters:

Model: frame_plus_pzn model -DEV-X1B
DB: EST -FASTA -LASTAP -SUFFIX -RST -MINMATCH=0.1 -LOWPCL=0 -LOPEXT=0
UNITS: bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.edi -LIST=45
-LOCAL=200 -THR_SCORE=pcv -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NUPM-ext -HEAPS-ext -MINLEN=0 -MAXLEN=200000000
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LOWPCL=0 -DEVP=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -XGAPOP=6 -XGAPEXT=7
YAPop 10 YGAPEXT=0.5 DELop=6 DELEXT=7

Database 1:

EST: *
1: em estba: *
2: em estbmc: *
3: em estlc: *
4: em estmq: *
5: em estov: *
6: em estpl: *
7: em estroz: *
8: em hte: *
9: qb est1: *
10: qb est2: *
11: qb hte: *
12: qb est3: *
13: qb est4: *
14: qb est5: *
15: em estbm: *
16: em estom: *
17: qb qss: *
18: em qss_hum: *
19: em qss_inv: *
20: em qss_pln: *
21: em qss_vit: *
22: em qss_fun: *
23: em qss_mam: *
24: em qss_mus: *
25: em qss_othe: *
26: em qss_pro: *
27: em qss_fat: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB	ID	Description
1	1551	38.6	1051	13	BM561083	BM561083 AGEN000R1
2	1448	36.0	881	14	BM220275	BM220275 AGEN000R1
3	1289	32.1	1001	13	BM475603	BM475603 AGEN000R1
4	1254	31.2	727	9	AL597811	AL597811 DKZP4130
5	1243	30.9	1364	21	AK011246	AK011246 Mus musc
6	1236.5	30.8	880	14	BM220275	BM220275 AGEN000R1
7	1208	30.1	835	13	BM453724	BM453724 AGEN000R1
8	1111	27.7	1223	13	BM450041	BM450041 AGEN000R1
9	1056.5	26.3	914	9	AA764411	AA764411 vw54a04.1
10	1043	26.0	752	9	AL564757	AL564757 AL564757
11	1037	25.8	605	9	AL599633	AL599633 DKZP4130
12	1030	25.6	694	10	BE082846	BE082846 R02 H1064
13	1016	25.3	623	10	AW583181	AW583181 JAL1109.7
14	1013	25.2	871	9	AA763685	AA763685 vw54a04.1
15	1034	25.0	922	12	BM253543	BM253543 BM253543
16	978	24.3	588	10	AW368234	AW368234 CM3 HT018
17	962	23.9	746	13	BI151568	BI151568 BM2916108
18	959	23.9	594	10	AV715533	AV715533 AV715533
19	955	23.8	584	12	BM149050	BM149050 BM08001.7
20	947.5	23.6	734	12	BM144857	BM144857 BM1790343
21	935	23.2	539	9	AL692790	AL692790 DKZP4131
22	938	22.6	797	10	BE33278	BE33278 BM1063967
23	882	22.0	522	9	AL120941	AL120941 LKZP27620
24	879.5	21.9	734	13	BI1526982	BI1526982 BM2927177
25	869.5	21.6	625	12	BM243088	BM243088 BM2358151
26	858	21.4	597	12	BM312436	BM312436 BM0-H1015
27	844	21.0	484	10	BE168095	BE168095 QV5-H1051
28	833.5	20.7	646	13	BM490416	BM490416 psp20.pk0
29	824	20.3	493	10	AW592887	AW592887 UT-HF BN0
30	815.5	20.3	522	9	AA072780	AA072780 BM76a05.1
31	784	19.5	2083	13	BM456728	BM456728 AGEN000R1
32	751	18.7	453	10	AW503514	AW503514 UT-HF BN0
33	684	17.0	438	12	BM27047	BM27047 R03 UN802
34	681.5	17.0	799	14	CM2304	CM2304 CM2304
35	667	16.6	651	13	BI1696841	BI1696841 BM3444967
36	659	16.4	480	10	AW951604	AW951604 EST363674
37	633	15.8	425	10	BE285796	BE285796 BM1096510
38	628	15.6	672	13	BI1868359	BI1868359 BM3492330
39	626	15.6	521	12	BM466536	BM466536 UT-HF BN0
40	625.5	15.6	487	12	BM998490	BM998490 CM2304
41	613	15.4	415	9	AT745644	AT745644 1124L00.X
42	611	15.2	410	12	BM560103	BM560103 UT-HF BN0
43	609	15.2	411	9	AT745643	AT745643 1124L00.X
44	593	14.8	580	9	AL599105	AL599105 DKZP4130
45	544.5	13.6	334	12	BM804092	BM804092 CM2304

ALIGNMENTS

RESULT	1
BM561083	1051 bp mRNA linear EST 20-FEB-2002
LOCUS	BM561083
DEFINITION	AGENCOURT_6566133 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:555905
ACCESSION	BM561083
VERSION	5.1 mRNA sequence
KEYWORDS	EST
SOURCE	human
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo;
TITLE	NIH MGC http://marc.ncbi.nlm.nih.gov/
	National Institutes of Health, Mammalian Gene Collection (MGC)

ORIGIN

Alignment Scores:
 Pred. No.: 5,746,146 Length: 881
 Score: 1448.00 Matches: 288
 Percent Similarity: 99.00% Conservative: 1
 Best Local Similarity: 97.96% Mismatches: 3
 Query Match: 96.05% Indels: 2
 DB: 14 Gaps: 0

US-09-026-459a-37 (1-781) x BM475603 (1-1001)

QY 61 MetGluAspSerValIleSerPheGluLeuMetGluCysValIleAspTyrPheIle 80
 DB 4 ATGGAAGATGATCTGGTGAATTCATTTGATGATGATGATGATGATGATGATGAT 62

QY 81 LysLeuSerProMetLeuLeuLysIleProTyrLysThrAlaValIleProLysAsn 100
 DB 63 AAACTCTCAGCTCCGATGCTCAAGGAAGCAATAAAGACAGCGTATACCCATTAA 122

QY 101 GlySerProArgThrProArgArgGlyGlnAsnArgSerAlaArgIleAlaLysGlnLeu 120
 DB 123 GGTTCACCTGGAACACCAAGGAGAGTGTAGAACAGAGTGTAGAACAGAGTGTAG 182

QY 121 GluAspSerPheArgIleIleGluValLeuCysLysGluHisGluCysAsnIleAspGlu 140
 DB 183 GAAATGATACAGGAATATATGAGTCTCTGTAAAGAACATGATGATATATAGATGAG 242

QY 141 ValLysAsnValTyrPheLysAsnPheIleProPheMetAsnSerLeuGlyLeuValThr 160
 DB 243 GTCAAAAATCTTATTCAAAAATTTTAACTTTTATCAATCTCTGGAGCTGTCAACA 302

QY 161 SerAsnGlyLeuProGlnValGlnAsnLeuSerLysArgTyrGluGlnIleTyrLeuLys 180
 DB 303 TCTAATGAGATCTGACAGGTGAAAATCTTCTAAACGATAGCAAGAAATTTATCTTAAA 362

QY 181 AsnLysAspLeuAspAlaArgIleGluPheLeuAspLysSerProGlnIleThrAspSer 200
 DB 423 ATACAGAGTTCACAAACACAGACACACACGAAAGATTAACCTTCAAGAGCTCAAT 482

QY 221 ValIleProGlnIleThrProValArgThrValMetAsnThrIleGluGlnLeuMetMet 240
 DB 483 GTAATTCCTGCAC 542

QY 241 IleLeuAsnSerAlaSerAspIleProSerGlnAsnIleSerLysPheAsnAspCys 260
 DB 543 ATTAAATTCACAGTCAATGACAGTTCACAGAGATGATGATGATGATGATGATGATG 602

QY 261 ThrValAsnProGlnIleSerIleLeuLysAspValLysAspIleGlyTyrIlePheLys 280
 DB 603 ATAGTCAATCCAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 662

QY 281 GlnGlnPheAlaLysValArgGlyGlnGlyCysValGlnIleGlySerGlnArgTyrLys 300
 DB 663 GAGAAATTCGTAAGACCTGAGAGAGGTTGTGTGAGAAATGGATATGATATGATATGA 722

QY 321 LeuGlyValArgIleGlyTyrArgValMetGluSerMetLeuLysSerGluGlnGlnArg 320
 DB 723 GTTGAATTCGTTGATATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 782

QY 341 LeuSerIleLeuAsnPheSerLysLeuLeuAsnAspAsnIlePheLysMetSerLeuL 340
 DB 783 TTAATTCATCAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 842

QY 361 GlnAlaCysAlaLeuGlnValValMetAlaThrTyrSer 362
 DB 843 TGAAGTGGCTCTTTGAGTTGTAATGAGGCAATATAGC 880

PAGE 3

BM475603

LOCUS
 DEFINITION BM475603 1001 bp mRNA linear EST 05 FEB 2002
 AGENE:JRT_6480907 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5575684
 5' mRNA sequence.
 ACCESSION BM475603
 VERSION
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 1001)
 AGINGERS NIH-MGC http://mgs.nhlbi.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rqb@ncs.nih.gov
 Tissue Procurement: APEC
 cDNA library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The L.M.A.G.E. Consortium (LML)
 DNA Sequencing by: Asenout Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the L.M.A.G.E. Consortium/LML at:
 http://image.llnl.gov
 Plate: LAM12326 Row: 1 Column: 04
 High quality sequence start: 17
 High quality sequence stop: 744.
 FEATURES
 Location/Qualifiers
 1..1001
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5575684"
 /clone_lib="NIH_MGC_92"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: testis; Vector: pCMV-Sp6R6; Site: 1; Note:
 Site: 2; S311; Cloned unidirectionally; oligo-dT primed;
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: This is a NIH-MGC Library."
 BASE COUNT 349 a 178 c 186 g 288 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,196,128 Length: 1001
 Score: 1289.00 Matches: 271
 Percent Similarity: 94.94% Conservative: 8
 Best Local Similarity: 91.25% Mismatches: 9
 Query Match: 92.09% Indels: 9
 DB: 14 Gaps: 1

US-09-026-459a-37 (1-781) x BM475603 (1-1001)

QY 1 MetSerArgLeuLeuLysIleTyrAspValLeuPheAlaLeuPheSerLysLeuGlnArg 20
 DB 77 ATCTCAAGACTGTTCAGGAAGATATCACTGATCTCTCTCTCTCTCTCTCTCTCT 146

QY 20 qThrCysGluLeuIleTyrLeuThrGlnProSerSerIleSerThrGlnIleAsnSer 40
 DB 137 GACACCTGGAACCTATAATAATTCACACCAACCAAGTACCTTCTCTCTCTCTCTCT 196

QY 40 rAlaLeuValLeuGlnGlnValSerIleIleIleIleIleIleIleIleIleIleIleIle 60
 DB 197 TGATTTGGTGGCTAAAAGTTTCTTGATGATACATTTTATATAGCTAAAGGGAAGATTA 256

QY 60 nMetGluAspAspLeuValIleSerPheGlnLeuMetLeuCysValIleLeuAspTyrPhe 80
 DB 257 AATGGAAGATGATCTGGTGAATTCATTTCACTTAATGCTATGCTGCTCTCTCTCT 316

QY 80 eLysLeuSerProProMetLeuLeuLysGlnProTyrLysThrAlaValIleProLysAs 100
 DB 317 TAAATCTCACCCTCCCACTGCTGCTCAAGCAATATATAAAACAGTCTGTAATATATA 376


```

104 914 TTTTGGATGCAATTAATAAATGAGGAGTGTGTTTGGAGGCTTTTAAAGCGGCGAG 857
209 TTTTGGATGCAATTAATAAATGAGGAGTGTGTTTGGAGGCTTTTAAAGCGGCGAG 226
106 856 AATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 798
229 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 248
106 797 TTTTGGATGCAATTAATAAATGAGGAGTGTGTTTGGAGGCTTTTAAAGCGGCGAG 736
249 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 268
106 747 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 679
269 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 288
106 678 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 623
289 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 308
106 622 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 563
309 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 328
106 562 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 503
329 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 348
106 502 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 443
349 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 368
106 442 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 383
369 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 388
106 382 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 323
389 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 408
106 322 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 263
409 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 428
106 262 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 203
429 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 448
106 202 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 146
449 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 468
106 145 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 86
469 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 488
106 85 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 26
489 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 496
106 25 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 2
RESULT 10
AL563757
DEFINITION AL563757.1 L11_NFL001_NH04 Homo sapiens cDNA clone c90n007A05.3
ACCESSION AL563757
VERSION AL563757.1 GI:12913464
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

```

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 792)
 Li W H, Gruber C, Jessee J, and Polayes D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope, Centre National de Séquençage
 BP 191 91006 Evry cedex France
 Email: secret.genoscope.cns.fr, Web: www.genoscope.cns.fr

FEATURES source

Location/Qualifiers
 1..792
 Jordanism="Homo sapiens"
 Zdb xref="taxon:9606"
 Zclone="CS000007A05"
 Zclone_lib="L11_NFL001_NH04"
 Zsex="male"
 Zissue_type="neuroblastoma cells"
 Zlab_host="pH10H"
 Znote="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with NotI and
 cloned into the NotI and EcoRV sites of the pCMVSPORT 6
 vector. Library was not radiolabeled. Library was constructed
 by life technologies. Contact: Feng Liang Life
 Technologies, a division of Invitrogen 9800 Medical Center
 Drive Rockville, Maryland 20850, USA Fax: (1) 401 610
 8471 Email: liang.feng@life-technologies.com URL:
<http://fulllength.invitrogen.com>

BASE COUNT ORIGIN

246 a 148 c 117 q 255 t 26 others
 Alignment Scores:
 Pred. No.: 4,71e-102 Length: 792
 Score: 1043.00 Matches: 217
 Percent Similarity: 84.53% Conservative: 7
 Best Local Similarity: 81.89% Mismatches: 38
 Query Match: 25.95% Indels: 3
 DB: 9 Gaps: 0
 US-09-026-459a-37 (1-781) x AL563757 (1-792)

```

QY 93 TTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 112
DB 791 AAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 732
QY 113 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 132
DB 731 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 672
QY 133 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 152
DB 671 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 612
QY 153 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 172
DB 611 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 552
QY 173 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 192
DB 551 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 492
QY 193 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 212
DB 491 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 452
QY 213 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 232
DB 431 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 372
QY 233 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 252
DB 371 ATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTTATGTTT 312

```



```

Db 424 TTAACAAAGGAGAGTAACCTTACACAGAGCAAAATGATAAAATGATGACGAC 265
QY 408 ISAPRILMetGluSerLeuAlaTyrLeuSerAspSerProLeuPheLeuLysG 428
Db 264 ATGAAATCATGAATCGTTGATGCTTTCAGATTCACGCTTATTGATCCATTAAAGC 205
QY 428 INSRYSAspAlaGluLysProThrAspHisLeuGluSerAlaCysProLeuAsnLeuP 448
Db 204 AGTAAAGATGAGAGAAAGAGCT--GATAAGCTTCAAGCTGCTGCTCAGAGCTGC 148
QY 448 roGluInAsnAspHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysL 468
Db 147 CTCTCAAGAGGAAACACATGAGACAAATAGATGCTTCCCTCAAGATCTCCAAAGCA 88
QY 468 ySLysGlySerThrArgValAspSerThrAlaAsnAlaGluThrAlaThrSerA 488
Db 87 AAAGAACTTCATCAATAGAGGAGAAATCTGCTGCAAAATACAGAGACAAAGAGCTGAG 28
QY 488 TApGluThrGluLysProLeuLys 496
Db 27 CATTTCATATCAATAGAGATTGAAA 2

PEPBLT 15
60254543
SEQUENCE
602464278F1 NIH MGC 90 Homo sapiens cDNA clone IMAGE:4472611 5'
mRNA sequence
60254543
60254543.1 51,1275,459
FSL
homosapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH MGC http://mgc.nhl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.,
Email: eap@nsl.nhl.nih.gov
Tissue Procurement: APC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequenced by: Incyte Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plated: L1AM10293 row: 1 column: 20
High quality sequence step: 643.
FEATURES
location/Qualifiers
1..922
/organism="Homo sapiens"
/zbb="Xenopus laevis"
/zclone="IMAGE:4472611"
/zclone_lib="NIH_MGC_90"
/zlab_host="pHILDB (phage-resistant)"
/znote="organ: liver; Vector: pCMV Sport6; Site1: NotI;
Site 2: SalI; cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full length clones and constructed by life technologies.
Note: this is a NIH_MGC library."
BASE COUNT 407 a 214 c 204 g 197 t
ORIGIN

```

```

Alignment Scores:
Prod. No.: 7,960,98 Length: 922
Score: 1004.00 Matches: 224
Percent Similarity: 90.51% Conservatve: 5
Best Local Similarity: 88.54% Mismatches: 21
Query Match: 24,99% Indels: 6
Gaps: 0

```

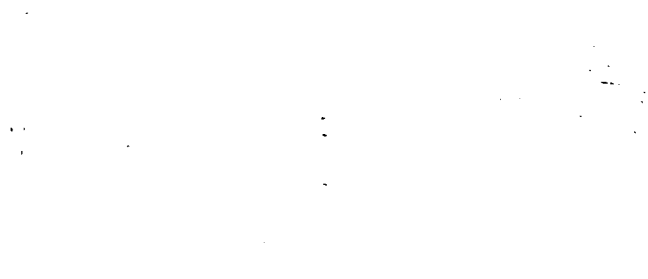
US-09-026-459a-37 (1 781) x 60254543 (1 922)

```

QY 532 TLeTTrPThrLeuPheThrHisThrLeuHisAsnGluTyrHisLeuMetLeuAspAr 551
Db 1 ATCATCTGGACGCTTTTCAGACACACGCTGCAGATGAGATGAGAACTCATAGACAG 60
QY 551 qHisLeuAspGlnHisMetMetCysSerMetTyrGlyTLeuCysLysValLysAsnHisAs 571
Db 61 GCATTG-GAGCAAAATATGATGATGCTGCAATGATGATGATATGATATGATATGATA 119
QY 571 pLeuLysPheLysLLeuValThrAlaTyrHisAspLeuProHisAlaValHisLeuThr 591
Db 120 CCTTAATTCAAAACAT-CTAACACACATACAAAGCACTTCCCACTGCTGCTGCTG 178
QY 591 rPheLysArgValLeuLLeuLysGluGluTyrAspSerLLeuValPheTyrAsnSer 611
Db 179 ATTCAAAAGCTGCT-CTGATCAAAACAAAGGAGATGATGATTTATTATAGATATTA 247
QY 611 rValPheMetGlnArgLeuLysThrAsnLLeuGluGlnTyrAlaSerThrArgProThr 631
Db 248 GGTCTTATGTCAGAGACTGAAGAAATAATTTGAGATAGCTTCAAGATGCTGCTG 296
QY 631 rLeuSerProLLeuProHisLLeuProArgSerProTyrLysPheProSerProLeuAr 651
Db 297 CTGTCACCAATACCTGACATACCTCAAGACCTCAACATTTCTAGTTACCTTACG 456
QY 651 qLeuProGlyGlyAsnLLeuTyrLLeuSerProLeuLysSerProTyrLysLLeuSerGlu 671
Db 457 GATTCTGTGAGGAGAAATATATATTTTCAATCTGAAACAAATATATATATATATAT 416
QY 671 yLeuProThrProThrLysMetThrProArgSerArgLLeuGluValSerLLeuLysLeu 691
Db 417 TCTGCAAAACAAACAAACAAACAAATGATCTGAAACAAATGATGATGATGATGAT 476
QY 691 rPheGlyThrSerGluLysPheGluLysLLeuAsnGluMetValLysAsnSerAspArgVa 711
Db 477 ATTGGGACCTTCACAGAGTCCAGAGAAAATAATCAGATGGTAGATGATGATGATG 536
QY 711 LLeuLysArgSerAlaGluGlySerAsnProProLysProLysProLysLysLeuArpHeAs 731
Db 537 GCTCAAAAGAGTGTGAGAGAGAGCAACGGTGTGTAATAAATCTGAAAAAATAGTTTGA 595
QY 731 pLeuGluGlySerAspGluAlaAsp-GlySerLysHis-LeuProGlyGluSerLysPhe 750
Db 596 TATTGAGAGATTCATGATGAAATGAGATGAGATGAGATGAGATGAGATGAGATGAG 655
QY 751 GlnGluLysLeuAlaGluMetThrSerThrArgMetGluGluSerLysGluMetLys 770
Db 656 CAGGCACAAAATGGGCAAAACGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 715
QY 771 AspSerMetAspThrSerAsnLysGluGluLys 781
Db 716 GGAAGCTGGGTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 746

```

Search completed: January 19, 2003, 05:06:10
Job time : 1739.9 secs



GenCure version 5.1.4
Copyright (c) 1993-2003 CompuGen Ltd.

com nucleotide complete search, using SW model

Run on: January 16, 2003, 15:20:17 : Search time 5160.97 Seconds
(without alignments)
17554.275 Million cell updates/sec

Filter: US-09-026-459A-36

Perfect score: 4114

Sequence: 1 GUGGACAGCAATATGATAGT 3113

Scoring table:

IDENTITY 90%

Gapop 10.0 : Gapext 1.0

Searches: 205440 seqs, 14531402878 residues

Total number of hits satisfying chosen parameters: 4106280

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: db_hum:*

2: db_hum:*

3: db_hum:*

4: db_hum:*

5: db_hum:*

6: db_hum:*

7: db_hum:*

8: db_hum:*

9: db_hum:*

10: db_hum:*

11: db_hum:*

12: db_hum:*

13: db_hum:*

14: db_hum:*

15: db_hum:*

16: db_hum:*

17: db_hum:*

18: db_hum:*

19: db_hum:*

20: db_hum:*

21: db_hum:*

22: db_hum:*

23: db_hum:*

24: db_hum:*

25: db_hum:*

26: db_hum:*

27: db_hum:*

28: db_hum:*

29: db_hum:*

30: db_hum:*

31: db_hum:*

32: db_hum:*

33: db_hum:*

34: db_hum:*

35: db_hum:*

36: db_hum:*

37: db_hum:*

38: db_hum:*

39: db_hum:*

40: db_hum:*

41: db_hum:*

Note: No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4107	99.8	4600	9	M4647 Human retin
2	4107	99.8	4859	6	AX44701 Sequence
3	4107	99.8	4849	9	L41970 Homo sapien
4	4105.4	99.8	3242	6	AK020341 Sequence
5	4105.4	99.8	3242	6	AK020342 Sequence
6	4105.4	99.8	3242	6	118496 Sequence 1
7	4105.4	99.8	3242	6	118497 Sequence 2
8	4105.4	99.8	4597	6	105411 Sequence 1
9	4105.4	99.8	4597	6	105411 Sequence 2
10	4099	99.6	4740	9	M15400 Human retin
11	4048.4	97.9	4580	6	A01444 H.sapiens D
12	2415	77.6	2994	6	AK091965 Sequence
13	2415	77.6	2994	6	AK091965 Sequence
14	2415	77.6	2994	6	AK091965 Sequence
15	2415	77.6	2994	6	AK091965 Sequence
16	2415	77.6	2994	6	AK091965 Sequence
17	2415	77.6	2994	6	AK091965 Sequence
18	2410.4	77.4	2995	6	AK144747 Sequence
19	2110.2	67.8	4591	6	E12560 cDNA encodi
20	2108.6	67.7	4591	10	M24391 Mouse retin
21	2098	67.4	4442	10	M24391 Mouse retin
22	1401.6	45.0	3456	5	U00114 Gallus gall
23	1394	44.8	1554	4	AF230742 Canis fam
24	1393.4	44.8	1554	5	AF230742 Canis fam
25	996.8	32.0	2808	5	AF230742 Canis fam
26	835	26.8	2352	6	109392 Sequence 26
27	835	26.8	2352	6	109392 Sequence 26
28	835	26.8	2352	6	109392 Sequence 26
29	835	26.8	2352	6	109392 Sequence 26
30	835	26.8	2352	6	109392 Sequence 26
31	835	26.8	2352	6	109392 Sequence 26
32	835	26.8	2352	6	109392 Sequence 26
33	835	26.8	2352	6	109392 Sequence 26
34	835	26.8	2352	6	109392 Sequence 26
35	835	26.8	2352	6	109392 Sequence 26
36	835	26.8	2352	6	109392 Sequence 26
37	835	26.8	2352	6	109392 Sequence 26
38	835	26.8	2352	6	109392 Sequence 26
39	835	26.8	2352	6	109392 Sequence 26
40	835	26.8	2352	6	109392 Sequence 26
41	835	26.8	2352	6	109392 Sequence 26

ALIGNMENTS

Result 1	HUMRA1RA	Human retinoblastoma associated (RH1) mRNA, complete cds.
LOCUS	HUMRA1RA	4600 bp Linear 12-JUL-1995
DEFINITION	Human retinoblastoma associated (RH1) mRNA, complete cds.	
ACCESSION	M4647.1	GI:190945
VERSION	M4647.1	GI:190945
KEYWORDS	retinoblastoma protein.	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens	
REFERENCE	1 (bases 1 to 4600)	
AUTHORS	Friend, S.H., Borowitz, J.M., Gerber, M.R., Ward, X.F., Beermann, E., Li, F.P., and Weisberg, R.A.	
TITLE	Deletions of a DNA sequence in retinoblastomas and mesenchyma	

QY 1507 TCACTGTTTTTAAAAAAGTATGAGGTATGCTATGCTGAGCTTAAATACACTTTGGAA 1566
 DB 1545 TCACTGTTTTTAAAAAAGTATGAGGTATGCTATGCTGAGCTTAAATACACTTTGGAA 2004
 QY 1567 CAGCTTTCTGCTGAGGAGCAGAAATACAAACATACATGAGCCTCTTCAGCAGACAC 1626
 DB 2005 CAGCTTTCTGCTGAGGAGCAGAAATACAAACATACATGAGCCTCTTCAGCAGACAC 2064
 QY 1627 CTGCAAAAGGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 1686
 DB 2065 CTGCAAAAGGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2124
 QY 1687 ATGATGAGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 1746
 DB 2125 ATGATGAGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2184
 QY 1747 TACAAAGGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 1806
 DB 2185 TACAAAGGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2244
 QY 1807 GAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 1866
 DB 2245 GAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2304
 QY 1867 ATTTGAGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 1926
 DB 2305 ATTTGAGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2364
 QY 1927 AGGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 1986
 DB 2465 AGGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2424
 QY 1987 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2046
 DB 2425 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2484
 QY 2047 AGATATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2106
 DB 2485 AGATATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2544
 QY 2107 AIAAATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2166
 DB 2545 AIAAATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2604
 QY 2167 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2226
 DB 2605 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2664
 QY 2227 AGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2286
 DB 2665 AGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2724
 QY 2287 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2346
 DB 2725 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2784
 QY 2347 AAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2406
 DB 2785 AAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2844
 QY 2407 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2466
 DB 2845 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2904
 QY 2467 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2526
 DB 2905 CAGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2964
 QY 2527 AGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2586
 DB 2965 AGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 3024
 QY 2587 AGTATGATATGAAATGATGAGCAGACAGGATTTGGCACTAAATATGAGTGTTC 2646

DB 3025 AGTATGAGCAGATGTTTCTCTCTCAAAATATAAAATGCTGCTGTTATGATGATAAAGA 4044
 QY 2647 ATGCTCTACAGTGGGAGTCTGATACCAACAGCAGCTGCTGCTGCTGCTGCTGCTG 2706
 DB 3085 ATGCTCTACAGTGGGAGTCTGATACCAACAGCAGCTGCTGCTGCTGCTGCTGCTG 3144
 QY 2707 TACATATAGGATGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2766
 DB 3145 TACATATAGGATGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3204
 QY 2767 AACATGAGCAGCTTACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2826
 DB 3205 AACATGAGCAGCTTACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3264
 QY 2827 CCCATTGAGCAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2886
 DB 3265 CCCATTGAGCAGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3324
 QY 2887 AAAATATGATATTTTACATGATGATTTTATTTTATTTTATTTTATTTTATTTTATTT 2946
 DB 3325 AAAATATGATATTTTACATGATGATTTTATTTTATTTTATTTTATTTTATTTTATTT 3384
 QY 2947 GCTGTTTTTATATAATTTTCTGCTTAAATATAAAGTGGAAATGATATATAATA 3006
 DB 3385 GCTGTTTTTATATAATTTTCTGCTTAAATATAAAGTGGAAATGATATATAATAATA 3444
 QY 3007 TCACTATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3066
 DB 3445 TCACTATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3504
 QY 3067 TATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3124
 DB 3505 TATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3564
 RESULT 2
 AX330701 Locus 4839 bp DNA Linear PAT 09 JAN 2002
 DEFINITION Sequence 1210 from Patent WO0194629.
 ACCESSION AX330701
 VERSION AX330701.1 GI:18104679
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1
 AUTHORS Young, P. E., Augustus, M., Carter, K. C., Ebner, R., Endress, G.,
 Horrikan, S., Soppet, D. K., and Weaver, Z.
 TITLE Cancer gene determination and therapeutic screening using signature
 gene sets
 JOURNAL Patent: WO 0194629-A 1210 13 08 2001;
 Avalon Pharmaceuticals (US)
 FEATURES
 Location/Qualifiers
 source 1..4839
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 HASH COUNT 1534 a 902 c 880 q 1523 t
 ORIGIN
 Query Match 99.8%; Score 3107; Ids 6; Length 4839;
 Host Local Similarity 100.0%; Prod. No. 0;
 Matches 3107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 7 ATGCTGAGCTGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 66
 DB 579 ATGCTGAGCTGTTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 638
 QY 67 ACACTGCACTTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 126
 DB 639 ACACTGCACTTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 698


```

GHDSPFWLWVNIKAFDPYKVIESIFKAGNLIRMIKHLERCRERIMESLAWLSD
SDPFLIKOSDRGPDHLESACPIELPIQNNHTAADMYLSVPSRKKKSGSTRVNS
TANATQTSFOTQKPLKSTSLFYKKVYRLAYLRLNLCERLLSEHPELEHIIWT
LEQHTQNEYELMRRLDQDMGSMYIGCKVKKNIDLKRIIVTAYKDIIPHAVQETFK
PLVKEEYDSIIIVYNSVEMQPLKTNILAYASTPPTSLPIHPSYKPESSPLR
IUGGNIYISPLKSPYKISEGLPTPTKMTIPSRHIVSICESFGISEKFKINQWVCNSD
PVIKPSAPGSDPKDKKIRHDIKGSDFANGSKHHPFSKPKQKLAHMTSTRTPMQKQ
KNDSMDISNKEER"
179
/ gene="RB1"
/ note="Tumor GOS16 unilateral mutation causing a tentative
wrong splice; G00-118-734"
/ citation=[5]
/ replace="t"
275..401
/ gene="RB1"
/ note="G00-118-734"
/ number=2
402..517
/ gene="RB1"
/ note="G00-118-734"
/ number=3
506..509
/ gene="RB1"
/ note="RBH64 bilateral familial deletion mutation causing
premature stop"
/ replace="ca"
518..637
/ gene="RB1"
/ note="G00-118-734"
/ number=4
547..542
/ gene="RB1"
/ note="Tumor GOS561 unilateral frameshift mutation causing
premature stop"
/ citation=[5]
/ replace="ca"
621..622
/ gene="RB1"
/ note="Tumor GOS37 bilateral frameshift mutation causing
premature stop; G00-118-734"
/ citation=[5]
/ replace="c"
638..676
/ gene="RB1"
/ note="G00-118-734"
/ number=5
677..744
/ gene="RB1"
/ note="G00-118-734"
/ number=6
745..855
/ gene="RB1"
/ note="G00-118-734"
/ number=7
856..944
/ gene="RB1"
/ note="G00-118-734"
/ number=8
999..1076
/ gene="RB1"
/ note="G00-118-734"
/ number=9
1077..1186
/ gene="RB1"
/ note="G00-118-734"
/ number=10
1095
/ gene="RB1"
/ note="Tumor GOS559 mutation causing premature stop ;
G00-118-734"
/ citation=[5]
/ replace="t"

```

```

variation
1104
/ gene="RB1"
/ note="Tumor GOS563 unilateral mutation causing premature
stop; G00-118-734"
/ citation=[5]
/ replace="t"
1187..1264
/ gene="RB1"
/ note="G00-118-734"
/ number=11
1209
/ gene="RB1"
/ note="Tumor GOS159 unilateral mutation causing a
premature stop; G00-118-734"
/ citation=[5]
/ replace="t"
1265..1352
/ gene="RB1"
/ note="G00-118-734"
/ number=12
1320
/ gene="RB1"
/ note="Tumor GOS13 unilateral frameshift mutation
resulting in premature stop; Associated with this mutation
is a large 13q14.3 sub-band deletion; G00-118-734"
/ citation=[5]
/ replace="ca"
1353..1469
/ gene="RB1"
/ note="G00-118-734"
/ number=13
1400..1401
variation
Query Match 99.8%; Score 3107; DB 9; Length 4859;
Best Local Similarity 100.0%; Pred. No. C;
Matches 3107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7 AUGTCAACACIGTGCAGACATGATGATTTGTTCACACCTTCAGCAAAATGCGAAGG 46
DB 579 ATGTCAAAGTTCACAAATGATGATTTGTTCACACCTTCAGCAAAATGCGAAGG 638
QY 67 ACATCGAACHTATATATTTGCACACACCCACACCTTCGATATCTATGCAAAATGAA 126
DB 639 ATATGTGAACHTATATATTTGCACACACCCACACCTTCGATATCTATGCAAAATGAA 698
QY 127 GTATTGCTGTAAAATTTCTTGAAATATATTTTATTATATGTAAGGGGAAGTATTACAA 186
DB 699 GATTTGGGCTTAAAAGCTTCTTCATATACATTTTATTAAGCTAAGGGGAAGTATTACAA 758
QY 197 AAGGAAACAGATGCTGCTGATTTGATTTGATTCAGTAAATGCTGCTGCTGCTGCTGCTGCT 246
DB 759 ATGTAAGATGATGCTGCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 816
QY 247 AAACCTCAGCTCCGACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406
DB 819 AAACCTCAGCTCCGACGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 874
QY 307 GGTTCATGCTGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 464
DB 879 GGTTCATGCTGAAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 948
QY 367 GAAATGATACAGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
DB 939 GAAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 998
QY 427 GCGAAAAATGCTTATTCAGCAAAATTTTATGCTTTTATGCTGCTGCTGCTGCTGCTGCTGCT 486

```



```

2647 ATGAGGCTTAAAGTGGGAGTCTCTATACACCGAGGCTCTCTGACTACTCTTGGCTTTTGG 2706
DB 4219 ATGAGGCTTAAAGTGGGAGTCTCTATACACCGAGGCTCTCTGACTACTCTTGGCTTTTGG 3278
QY 2207 TACGATATAGGAGTGGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2766
DB 4279 TACGATATAGGAGTGGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3338
QY 2767 AACATCAATACCGCTTACGAAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2826
DB 4339 AACATCAATACCGCTTACGAAATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3398
QY 2827 GCAATTCACCAAAATATGCTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2886
DB 4399 GCAATTCACCAAAATATGCTGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3458
QY 2887 AAAAAATACATATACATATACATATATATATATATATATATATATATATATATATATATAT 2946
DB 4459 AAAAAATACATATACATATATATATATATATATATATATATATATATATATATATATATAT 3518
QY 2947 GCTTGTGTTTAAATATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3006
DB 4519 GCTTGTGTTTAAATATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3578
QY 3007 TGAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3066
DB 4579 TGAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 3638
QY 3647 TATTTTCTGATGCACTATGCTTTTAAATGAGGATTTATGATAGT 3113
DB 4649 TATTTTCTGATGCACTATGCTTTTAAATGAGGATTTATGATAGT 3695

RESULTS: 4
AR072041 3232 bp DNA linear PAT 18 FEB 2000
DEFINITION Sequence 1 from patent US 5912236.
ACCESSION AR072041
VERSION AR072041.1 GI:7222919
KEYWORDS tumor suppressor gene therapy
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. Hudes 1 to 3232
AUTHORS Xu, H.-J.; Hu, S.-X. and Benedict, W.F.
TITLE Broad-spectrum tumor suppressor genes: gene products and methods for
tumor suppressor gene therapy
JOURNAL Patent: US 5912236-A 15-JUN-1999.
FEATURES
Location/Qualifiers
Source 1..3232
organism="Unknown"
BASE COUNT 1086 a 597 c 566 g 983 t
GC 51%
Query Match 99.8%; Score 3105.4; DB 6; Length 3232;
Best Local Similarity 100.0%; E-Val. No. 0;
Matches 3106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 AATGCAAGCTGTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 66
DB 124 ATGCAAGCTGTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
QY 67 AATGCAAGCTGTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 126
DB 184 AATGCAAGCTGTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 243
QY 127 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 186
DB 244 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 303
QY 187 AATGCAAGCTGTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
DB 404 AATGCAAGCTGTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 363

```

```

QY 247 AAATCTCAGCTCCCATGTTGCTCAAGAGCATATAAAACAGCTGTTATACCAATTAAT 406
DB 364 AAATCTCAGCTCCCATGTTGCTCAAGAGCATATAAAACAGCTGTTATACCAATTAAT 423
QY 307 GGTTCATCTTAAATATATATATATATATATATATATATATATATATATATATATATAT 466
DB 424 GGTTCATCTTAAATATATATATATATATATATATATATATATATATATATATATATAT 483
QY 367 GAAAAATCATAAAGAAATATATATATATATATATATATATATATATATATATATATATAT 426
DB 484 GAAAAATCATAAAGAAATATATATATATATATATATATATATATATATATATATATATAT 543
QY 427 GTGAAAAATGTTTATTTTCAAAAAATTTTATACCTTTTATGAAATCTCTGACCTGTAACA 486
DB 544 GTGAAAAATGTTTATTTTCAAAAAATTTTATACCTTTTATGAAATCTCTGACCTGTAACA 603
QY 487 TCTATGAACTTCCAGAGTTGAAAAATCTTTCTAACGATAGAGAGAAATTTATCTTAAA 546
DB 604 TCTATGAACTTCCAGAGTTGAAAAATCTTTCTAACGATAGAGAGAAATTTATCTTAAA 663
QY 547 AATAAAGATCTAGATGCAAGATTTATTTTGGATCATATAAAACCTCTGACATGATGAT 606
DB 664 AATAAAGATCTAGATGCAAGATTTATTTTGGATCATATAAAACCTCTGACATGATGAT 723
QY 607 ATAGAATCTTTTGAATCAATACAGAAATATATATATATATATATATATATATATATATAT 666
DB 724 ATAGAATCTTTTGAATCAATACAGAAATATATATATATATATATATATATATATATATAT 783
QY 667 GTAATCTCTTACATCTCTGAGTTTAAAGATTTATATATATATATATATATATATATATAT 726
DB 784 GTAATCTCTTACATCTCTGAGTTTAAAGATTTATATATATATATATATATATATATATAT 843
QY 727 ATTTTAAATCTCAGAGCTCAACCTTCAGAAAATCTGATTTCTCTATTTTAAACAACTG 786
DB 844 ATTTTAAATCTCAGAGCTCAACCTTCAGAAAATCTGATTTCTCTATTTTAAACAACTG 903
QY 787 AGATGCTCAATCTTAAAGAGTATATATATATATATATATATATATATATATATATATATAT 846
DB 904 AGATGCTCAATCTTAAAGAGTATATATATATATATATATATATATATATATATATATATAT 963
QY 847 GAGAAATTTGCTTAAAGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 906
DB 964 GAGAAATTTGCTTAAAGCTGTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCT 1023
QY 907 GTTGGAGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 968
DB 1024 GTTGGAGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1083
QY 967 TTATGCTTCAAAATTTTAAAGAGTATATATATATATATATATATATATATATATATATAT 1026
DB 1084 TTATGCTTCAAAATTTTAAAGAGTATATATATATATATATATATATATATATATATATATAT 1143
QY 1027 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1086
DB 1144 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
QY 1087 TGTGAAACAGATGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1146
DB 1204 TGTGAAACAGATGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1263
QY 1147 TTTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
DB 1264 TTTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
QY 1267 AAATCTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1266
DB 1324 AAATCTTAAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
QY 1267 CCTTTATTTGATCTTTTAAACATCAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1326
DB 1384 CCTTTATTTGATCTTTTAAACATCAAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCT 1443

```


11 1064 CAGAAATTTCTAAAGCTGTGGGAAAGGTTGTGTGTGAAAAATGGATACAGAGATACAAA 1023
QY 1067 CTTGGACCTGGCTTGATTAACGAGTAATACGAATCCATGCTTAATACAGACAGACACAGA 966
DB 1024 CTTGGACCTGGCTTGATTAACGAGTAATACGAATCCATGCTTAATACAGACAGACACAGA 1083
QY 1067 TTATGCAATCAMAATTTTAGCAAACTCTGTAATGATGACAGCATTTTTCATATGCTTTATG 1026
DB 1084 TTAACCAATCAAAAATTACCAAACTCTGTAATGATGACAGCATTTTTCATATGCTTTATG 1143
QY 1027 GGTGGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAG 1086
DB 1144 GGTGGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAG 1203
QY 1087 TCTGGAAACAGATTTGCTTTCCATGGAATCTGGAATGCTGTAATTTAAAGGCTTTTCAT 1146
DB 1204 TCTGGAAACAGATTTGCTTTCCATGGAATCTGGAATGCTGTAATTTAAAGGCTTTTCAT 1263
QY 1147 TTTTAAATATGATGTAATTTTATCAAGGTAAGGTAATTTTAAAGGTAATTTTAAAGGTA 1206
DB 1264 TTTTAAATATGATGTAATTTTATCAAGGTAAGGTAATTTTAAAGGTAATTTTAAAGGTA 1323
QY 1267 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
DB 1324 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1383
QY 1267 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1326
DB 1384 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1443
QY 1327 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386
DB 1444 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1503
QY 1387 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
DB 1504 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1563
QY 1447 GCAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 1506
DB 1564 GCAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 1623
QY 1567 TCAATGCTTTTATAAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1566
DB 1624 TCAATGCTTTTATAAAAAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1683
QY 1567 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1626
DB 1684 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1743
QY 1627 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1686
DB 1744 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1803
QY 1787 ATGATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1746
DB 1804 ATGATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1863
QY 1747 TATAAGGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806
DB 1864 TATAAGGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
QY 1867 GAGTATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1866
DB 1924 GAGTATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1983
QY 1867 ATTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1926
DB 1984 ATTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2043
QY 1927 AGAGCTTACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1986
DB 2044 AGAGCTTACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103

QY 1987 GCGCTGAGAGTCCATATATAAATTTTCAAGAAAGTCTGCCAAATACCAATCAAAATTAATCTCA 2046
DB 2104 GCGCTGAGAGTCCATATATAAATTTTCAAGAAAGTCTGCCAAATACCAATCAAAATTAATCTCA 2163
QY 2047 ATATCAAAATGCTTATATCAATTTGCTGAAATCAATTTGCTGAAATTTGCTGAAATTTGCTG 2106
DB 2164 ATATCAAAATGCTTATATCAATTTGCTGAAATCAATTTGCTGAAATTTGCTGAAATTTGCTG 2223
QY 2107 ATAAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2166
DB 2224 ATAAATCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2283
QY 2167 GCTGCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226
DB 2284 GCTGCTGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2343
QY 2227 ATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2286
DB 2344 ATAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2403
QY 2287 GCAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 2346
DB 2404 GCAAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGAC 2463
QY 2347 AAATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2406
DB 2464 AAATGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2523
QY 2407 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2466
DB 2524 GAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2583
QY 2467 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2526
DB 2584 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2643
QY 2527 AGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2586
DB 2644 AGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2703
QY 2587 AGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2646
DB 2704 AGTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2763
QY 2647 ATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2706
DB 2764 ATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2823
QY 2707 TAGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2766
DB 2824 TAGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2883
QY 2767 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2826
DB 2884 AGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2943
QY 2827 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2886
DB 2944 GCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3003
QY 2887 AAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2946
DB 3004 AAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3063
QY 2947 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3006
DB 3064 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3123
QY 3007 TGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3066
DB 3124 TGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3183

QY 306/ TATTTTTCATGCAATTTATGTTTATTAATGAGGATTTATGATAGT 4114
 DB 3184 TATTTTTCATGCAATTTATGTTTATTAATGAGGATTTATGATAGT 4230

RESULT 7
 11847/6
 DEFINITION Sequence 2 from patient US 5496741.
 11847/
 VERSION 11847.1 GI:1598852
 KEYWORDS
 SOURCE Unknown.
 ORGANISM
 Unclassified.
 REFERENCE 1 (bases 1 to 4242)
 AUTHORS Xu, H., J., Hu, S., X., and Renodiet, W. F.
 TITLE Broad spectrum tumor suppressor genes, gene products and methods
 for tumor suppressor gene therapy
 JOURNAL Patent US 5496741 A 2 05-MAR-1996
 FEATURES
 Location/Qualifiers
 1..4242
 /organism="unknown"

BASE COUNT 983 a 566 c 597 g 1086 t
 GC101N

Query Match 99.9%; Score 4105.4; DB 6; Length 4242;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 109; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 ATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 66
 DB 3113 ATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 3054

QY 67 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 126
 DB 3054 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2994

QY 127 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 186
 DB 2994 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2944

QY 187 ATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 246
 DB 2944 ATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2874

QY 247 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 306
 DB 2874 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2814

QY 307 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 366
 DB 2814 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2754

QY 367 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 426
 DB 2754 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2694

QY 427 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 486
 DB 2694 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2634

QY 487 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 546
 DB 2634 GATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2574

QY 547 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 606
 DB 2574 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2514

QY 607 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 666
 DB 2514 AATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2454

QY 667 GAAATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 726
 DB 2453 GAAATGTCAGACATGTCGAGAAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2494

QY 727 ATTTTAAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 786
 DB 2493 ATTTTAAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2444

QY 787 ACATGCAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 846
 DB 2444 ACATGCAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2274

QY 847 GAGCAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 906
 DB 2273 GAGCAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2214

QY 907 CTTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 966
 DB 2213 CTTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2154

QY 967 TATTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1026
 DB 2153 TATTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2094

QY 1027 GATTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1086
 DB 2093 GATTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 2044

QY 1087 TCTTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1146
 DB 2043 TCTTCGATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1974

QY 1147 TTTTACAAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1206
 DB 1973 TTTTACAAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1914

QY 1207 AATCATTTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1266
 DB 1913 AATCATTTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1854

QY 1267 CTTTATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1326
 DB 1853 CTTTATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1794

QY 1327 TCTGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1386
 DB 1793 TCTGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1744

QY 1387 TCTGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1446
 DB 1743 TCTGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1674

QY 1447 GATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1506
 DB 1673 GATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1614

QY 1507 TATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1566
 DB 1613 TATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1594

QY 1567 GATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1626
 DB 1553 GATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1494

QY 1627 CTGCAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1686
 DB 1493 CTGCAATTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1444

QY 1687 ATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1746
 DB 1443 ATGTCAGCAATGATGATGTTTGGTCTTTCAGCAATTTGGAAGG 1474


```
247 ATGGCTGACAGTGGAGTGTGATAGGAGGAGGAGTGTGATGTTGGCTTTG 2706
|||||
682 AAGCTCTACCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 3141
|||||
307 GGTTCATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2766
|||||
742 GGTTCATGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3201
|||||
367 GAAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2826
|||||
802 GAAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3261
|||||
427 GTGAAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2886
|||||
862 GTGAAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3321
|||||
487 TCTAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2946
|||||
922 TCTAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3381
|||||
547 AATTAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3006
|||||
982 AATTAATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3441
|||||
607 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3066
|||||
1042 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3501
|||||
667 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3113
|||||
1102 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 3548
|||||
727 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 786
|||||
1162 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1221
|||||
787 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 846
|||||
1222 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1281
|||||
847 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 906
|||||
1282 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1341
|||||
907 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 966
|||||
1342 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1401
|||||
967 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1026
|||||
1402 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1461
|||||
1027 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1086
|||||
1462 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1521
|||||
1087 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1146
|||||
1522 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1581
|||||
1147 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1206
|||||
1582 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1641
|||||
1207 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1266
|||||
1642 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1701
|||||
1267 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1326
|||||
1702 ATAGATGCAACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1761
|||||
```

```
RESULT 1
109369 109369 4597 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent. WO 8906703.
ACCESSION 109369
VERSION 109369.1 GI:587899
KEYWORDS
SOURCE unknown.
ORGANISM unknown.
REFERENCE 1 (bases 1 to 4597)
AUTHORS Bryant, P., Friend, S., and Vandell, D.W.
TITLE DIAGNOSIS OF RETINOBLASTOMA
JOURNAL Patent: WO 8906703-A 1 27 JUL 1989
FEATURES
Location/Qualifiers
SOURCE 1..4597
RASP (SDN) 1489 a 841 c 813 g 1454 t
ORIGIN
```

```
Query Match 99.8%; Score 3105.4; DB 6; Length 4597;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 AAGTCAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 66
|||||
L4 442 AAGTCAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 501
|||||
QY 67 AAGTCAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 126
|||||
L6 502 AAGTCAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 561
|||||
QY 127 GATGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 186
|||||
L8 562 GATGAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 621
|||||
QY 187 AAGTCAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 246
|||||
L9 422 AAGTCAAGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 681
|||||
```


27 1427 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1486
 146 1762 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1821
 27 1487 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1446
 146 1842 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1881
 27 1447 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1506
 146 1882 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1941
 27 1507 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1566
 146 1942 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2001
 27 1567 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1626
 146 2002 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2061
 27 1627 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1686
 146 2062 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2121
 27 1687 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1746
 146 2122 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2181
 27 1747 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1806
 146 2182 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2241
 27 1807 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1866
 146 2242 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2301
 27 1867 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1926
 146 2302 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2361
 27 1927 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 1986
 146 2362 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2421
 27 1987 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2046
 146 2422 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2481
 27 2047 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2106
 146 2482 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2541
 27 2107 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2166
 146 2542 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2601
 27 2167 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2226
 146 2602 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2661
 27 2227 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2286
 146 2662 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2721
 27 2287 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2346
 146 2722 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2781
 27 2347 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2406
 146 2782 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2841
 27 2407 TTTGTTTGGCTTTAAATTTTGGTCTGAGAAATATATACACTGACACATATATATATCT 2466

2842 CAGATGAGCTGATATACATTTCCAGGCTTCTGTTTATGGGCAATATATATATCTTCTAGCT 2901
 2467 CTTTCTGTCGAT 2526
 2462 CTTTCTGTCGAT 2491
 2527 ATGTAT 2586
 2462 ATGTAT 2521
 2587 ATGTAT 2646
 4022 AGTTGATGAGAT 4081
 2647 ATGTCCTTACAGTCGAGAT 2706
 4082 ATGTCCTTACAGTCGAGAT 4141
 2707 TAGAT 2766
 4142 TAGAT 4201
 2767 ATATGAT 2826
 3462 ATATGAT 3521
 2827 CCATTCACCAAAAT 2886
 3522 CCATTCACCAAAAT 3581
 2887 AT 2946
 3422 AT 3481
 2947 GCTTGTAT 3006
 3482 GCTTGTAT 3541
 3007 TGAT 3066
 3442 TGAT 3501
 3067 TATTTCTTAT 3126
 3502 TATTTCTTAT 3561
 RESULT 10
 HUMRS
 LOCUS Human retinoblastoma susceptibility mRNA, complete cds. 4740 bp mRNA linear DEC 12 2001 1995
 DEFINITION M15400
 ACCESSION M15400.1 GI:190958
 VERSION retinoblastoma susceptibility.
 KEYWORDS Homo sapiens (clone: RB [1.5]) testis retina cDNA to mRNA.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 REFERENCES Chordata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo
 1 (bases 243 to 4740)
 Lee, W.H., Bookstein, R., Hong, F., Young, L.J., Shew, J.Y., and Lee, E.Y.
 Human retinoblastoma susceptibility gene: cloning, identification,
 and sequence
 Science 235 (4794), 1494-1499 (1987)
 JOURNAL B7149066
 MEDLINE 3823889
 PUBMED 2 (bases 1 to 480)
 REFERENCE Lee, W.H., Bookstein, R., Young, L.J., Lin, C.J., Rosenfeld, M.G., and
 Lee, W.H.
 Molecular mechanism of retinoblastoma gene inactivation in
 retinoblastoma cell line Y79
 Proc. Natl. Acad. Sci. U.S.A. 85 (16), 6017-6021 (1988)
 JOURNAL B8320373
 MEDLINE


```

14 1440 TTATGTAATGAAAAATTTTACCAAAATCTGGAATGACCAACATTTTTCATATGCTTTTATG 1599
15 1441 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1086
16 1442 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1659
17 1443 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1146
18 1444 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1719
19 1445 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1206
20 1446 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1779
21 1447 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1266
22 1448 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1839
23 1449 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1326
24 1450 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1899
25 1451 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1386
26 1452 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1959
27 1453 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1446
28 1454 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2019
29 1455 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1506
30 1456 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2079
31 1457 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1566
32 1458 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2139
33 1459 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1626
34 1460 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2199
35 1461 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1686
36 1462 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2259
37 1463 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1746
38 1464 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2319
39 1465 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1806
40 1466 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2379
41 1467 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1866
42 1468 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2439
43 1469 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1926
44 1470 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2499
45 1471 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 1986
46 1472 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2559
47 1473 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2046
48 1474 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2619
49 1475 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2106
50 1476 GGTGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 2679

```

```

QY 2107 ATAAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2166
DB 2680 ATAAATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2749
QY 2167 GTCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
DB 2740 GTCCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2799
QY 2227 AGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2286
DB 2800 AGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2859
QY 2787 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2846
DB 2860 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2919
QY 2347 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2406
DB 2920 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2979
QY 2407 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2421
DB 2980 CACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2994

RESULT 13
AR098189
LOCUS 2994 bp DNA linear PAT 14-FEB-2003
DEFINITION Sequence 3 from patent US 6074850.
ACCESSION AR098189
VERSION AR098189.1 GI:12807446
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Antelman,D., Gregory,P.J. and Wills,K.N.
TITLE Pct:inhibitors fusion polypeptides
JOURNAL Patent: US 6074850-A 3 13-JUN-2000;
FEATHEPS Location/Qualifiers
SOURCE 1. 2994
BASE COUNT 974 a 618 c 593 g 809 t
ORIGIN
Query Match 77.6%; Score 2415; Pos 6; Length 2994;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2415; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGTCGACACTGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 65
DB 580 AGTCGACACTGTCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 639
QY 67 AATGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 126
DB 640 AATGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 699
QY 107 GATGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 186
DB 700 GATGAGGAGTCTTGAGGTTGTAATGAGGAGATATAGCAGAGTACATGACAGATCTTCAT 719
QY 187 AGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 246
DB 760 AGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 819
QY 247 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 406
DB 820 AAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879
QY 307 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 366
DB 630 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649

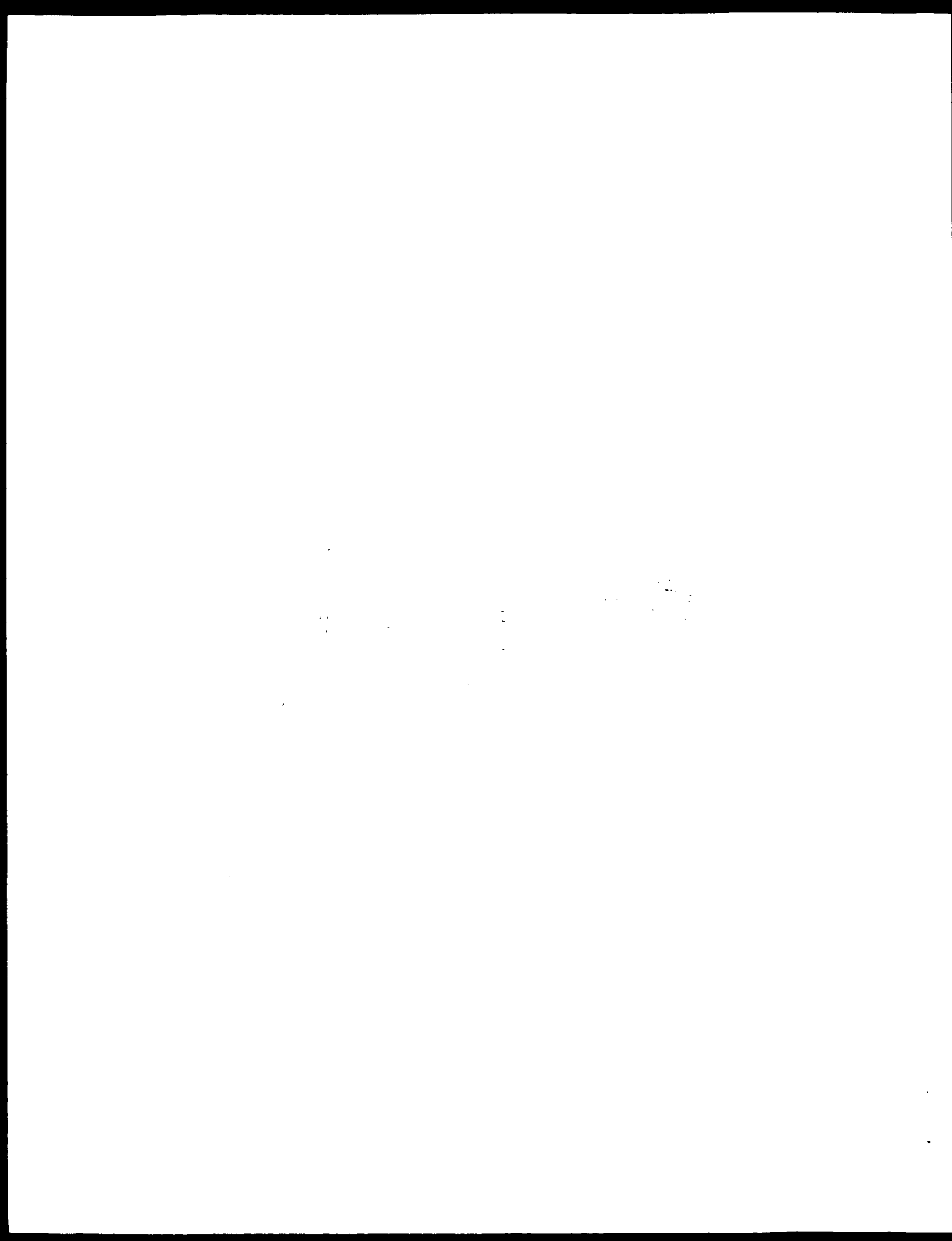
```


[illegible]

AR091945	AR091945	2994 bp	DNA	linear	PAT 08-SEP-2000
DEFINITION	Sequence 4 from patent US 5998134.				
ACCESSION	AR091945				
VERSION	AR091945.1	GI-10019719			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 2994)				
AUTHORS	Lee, W. H. and Lee, E. Y.-H. P.				
TITLE	Retinoblastoma gene-cancer suppressor and regulator				
JOURNAL	US 5998134 A 27 SEP 1999				
FEATURES	Location/Qualifiers				
1..2994					
/organism="unknown"					
BASE COUNT	974 a 618 c 594 g 808 t				
ORIGIN					
Alignment Scores:	4 462 294	Length:	2994		
Prod. No.	4017 00	Matches:	781		
Score	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
DB:	6	Gaps:	0		
US 09-026 459A-47 (1-781) x AR091945 (1-2994)					
QY	1 MetSerArgLeuLeuLeuLysLysTyrAspValLeuPheAlaLeuPheSerLysLeuGluArg	20			
DB	580 ATGCAAGACGCTTTGAAGAACTATGATGATTTGTTGACCTTTGSCAAATGSAAGG	639			
QY	21 ThrCysGluLeuLeuLeuLeuThrGlnProSerSerSerSerSerSerSerSerSerSer	40			
DB	440 AATGTAATATATATATATGACAAACCCAGCAGTTCATATCTATGTAATATATCT	699			
QY	41 AlaLeuValLeuLysValSerTrpIleThrPheLeuLeuAlaLysGlyGluValLeuGln	60			
DB	700 GCAATGGCGCTAAAGATTTCTTGGATCACAATTTTATAGCTAAAGGGGAAGATTCAA	759			
QY	61 MetGluAspLeuValIleSerPheGlnLeuMetLeuCysValLeuAspTyrPheIle	80			
DB	760 ATGAAACATGATCTGATGATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT	819			
QY	81 LysLeuSerProProMetLeuLeuLysGluProTyrLysThrAlaValIleProIleAsn	100			
DB	820 AAAGCTCTCACTCCATGTTGCTCAAGAACCATATATAAACAGCTGTATACCCATTAT	879			
QY	101 GlySerProArgThrProArgArgGlnAsnArgSerAlaAlaGluIleAlaLysGluLeu	120			
DB	880 GGTTCACCTCGAACATCCAGCGGAGGTCAGAACAGCAGTGCAGGATAGCAAAACAATA	949			
QY	121 GluAspAspThrArgIleIleGlnValLeuCysLysGluHisGluCysAsnIleAspGlu	140			
DB	940 GAAATGATACAGCAATTTATTAATCTCTCTGTAAGAACATGATTAATATATAGATGAG	999			
QY	141 ValLysAsnValIrrPheLysAsnIleIleProPheMetAsnSerLeuGlyLeuValThr	160			
DB	1000 GTGAAAATGTTTATTTCAAAAATTTTATACCTTTTATGAAATCTCTGAGACITGTACA	1059			
QY	161 SerAsnGlyLeuProGluValGluAsnLeuSerLysArgTyrGluGluIleTyrLeuLys	180			
DB	1060 GCTAACTGACCTTCCACAGCTTGAAATCTTCTAAACCATACGAGAAATTTATCTTAAA	1119			
QY	181 AsnLysAspLeuAspAlaArgLeuPheLeuAspHisAspLysThrLeuGlnThrAspSer	200			
DB	1120 AATAAAGATCTAGATGCAAGATTTATTTTGGATCATGATATAAAGCTCTTCAGACTGATCT	1179			
QY	201 ThrAspSerPheCysThrGlnArgThrProArgLysSerAsnLysAspGluValAsn	220			
DB	1180 ATGACAGTTTTGAAACACAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	1239			
QY	221 ValIleProProHisThrProValArgThrValMetAsnThrIleGlnLeuMetMet	240			
DB	1240 GTAATTCCTCCACACATCCAGCTTACGACGCTTATGACACACATATCCACAAATATATGATC	1299			
QY	241 ThrLeuAsnSerAlaSerAspCysProSerGluAsnLeuIleCysTyrPheAsnAspCys	260			
DB	1300 ATTTTAAATTCAGCAAGGATCAAAATTCAGAAAAATGATTTCTATTTTAAVAACTG	1359			
QY	261 ThrValAsnProLysCysSerIleLeuLeuLysArgValLysAspIleCysTyrIlePheLys	280			
DB	1360 AATGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCA	1419			
QY	281 GluLysPheAlaLysAlaValGlyGlyLysValGluIleGlySerGluArgTyrLys	300			
DB	1420 GACAAATTTGCTAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1479			
QY	301 LeuGlyValArgLeuTyrTyrArgValMetGluSerMetLeuLysSerGluGluArg	320			
DB	1480 CTTGGAGTTGCTTCTATTTATGAGTAAATGGAATGGAATGGAATGGAATGGAATGGAATG	1539			
QY	321 LeuSerIleCysAsnIleSerLysLeuLeuAspAspAspIleLeuHisMetSerLeuLeu	340			
DB	1540 TATCTATTTAAAAATTTAGTAAATCTGCAATGCAAAATTTTATGATGCTTTATG	1599			
QY	341 AlaCysAlaLeuGluValValMetAlaTrpTyrSerArgSerThrSerGlnAsnLeuAsp	360			
DB	1600 GGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1659			
QY	361 SerGlyThrAspLeuSerPheProTrpIleLeuAsnValLeuAsnLeuLysAlaPheAsp	380			
DB	1660 TTTGAAATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1719			
QY	381 PheTyrLysValIleCysSerPheIleLysAlaGluGlyAsnLeuThrArgGluMetIle	400			
DB	1720 TTTTCAAGTATGATGCAAGTCTTATCAAGGACAAAGCAAGCAAGCAAGCAAGCAAGCA	1779			
QY	401 LysHisLeuGluArgCysGluHisArgIleMetGluSerLeuAlaTrpLeuSerAspSer	420			
DB	1780 AAACATTTAGACGATGTCACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA	1839			
QY	421 ProLeuPheAspLeuIleLysGlnSerLysAspArgGluGlyProThrAspHisLeuGlu	440			
DB	1840 CTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT	1899			
QY	441 SerAlaCysProLeuAsnLeuProLeuGlnAsnAsnHisIleThrAlaAlaAspMetTyrLeu	460			
DB	1900 TCTGCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1959			
QY	461 SerProValArgSerProLysLysLysSerThrThrArgValAsnSerThrAlaAsn	480			
DB	1960 TCTCTGTAAGATCTGTAAGATCTGTAAGATCTGTAAGATCTGTAAGATCTGTAAGATCTG	2019			
QY	481 AlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLysProLeuLysSerThrSerLeu	500			
DB	2020 GAT	2079			
QY	501 SerLeuPheTyrLysLysValTyrArgLeuAlaTyrIleGluAsnThrLeuCysGlu	520			
DB	2080 TATGCTTTTATATAAAATGTAATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT	2139			
QY	521 ArgLeuLeuSerGluHisProGlnLeuGluHisIleLeuIleThrLeuPheGlnHisThr	540			
DB	2140 GCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2199			
QY	541 LeuGlnAsnGluTyrGlnLeuMetArgAspArgHisLeuAspGlnIleLeuMetCysSer	560			
DB	2200 TTTTAAATGAT	2259			
QY	561 MetTyrGlyIleCysLysValLysAsnIleAspLeuLysPheLysIleIleValThrAla	580			
DB	2260 ATGATGCTGAT	2319			
QY	581 TyrLysAspLeuProHisAlaValGlnGluThrPheLysArgValLeuIleLysGluGlu	600			

[illegible]

[illegible]



Genome version 5.1.3
Copyright (c) 1995-2003 CompuGen Ltd.

cm nucleic nucleic search, using sw model

Run on: January 16, 2003, 15:20:17 : Search time 5509.12 Seconds
(without alignments)
17594.275 Million cell updates/sec

File: us_09_026_459a_32

Perfect score: 3323
Sequence: 1 GCGAATGAGGAGAAAGTTTC.....AAATGAGATTATGATAG 3323

Scoring table:
IDENTITY_NDC
Gapop 10.0 : Gapext 1.0

Searched: 2054630 seqs, 1455140279 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum hit seq length: 0

Maximum hit seq length: 2000000000

Post processing: Minimum Match 100%

Listing first 45 summaries

Database: GenBank

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_cm:*

5: gb_ov:*

6: gb_pa:*

7: gb_ph:*

8: gb_pi:*

9: gb_pr:*

10: gb_ro:*

11: gb_st:*

12: gb_sy:*

13: gb_un:*

14: gb_vt:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_cm:*

21: em_or:*

22: em_ov:*

23: em_pa:*

24: em_ph:*

25: em_pi:*

26: em_ro:*

27: em_st:*

28: em_un:*

29: em_vt:*

30: em_hgt_hum:*

31: em_hgt_inv:*

32: em_hgt_of_bst:*

33: em_hgt_mus:*

34: em_hgt_pla:*

35: em_hgt_rtd:*

36: em_hgt_mam:*

37: em_hgt_vit:*

38: em_sy:*

39: em_hgt_hum:*

40: em_hgt_mus:*

41: em_hgt_ofther:*

Prod. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	3315	99.8	4600	9	HUMHRAIRA
2	3315	99.8	4600	9	AX336703
3	3315	99.8	4600	9	AX336703
4	3313.4	99.7	4597	6	HUMHRAIRA
5	3313.4	99.7	4597	6	HUMHRAIRA
6	3280.8	98.7	4740	9	HUMHRAIRA
7	3254.8	97.9	4580	6	A01444
8	3222.4	97.0	4580	6	AK037303
9	3222.4	97.0	4580	6	AK037303
10	3222.4	97.0	4580	6	AK037303
11	3222.4	97.0	4580	6	AK037303
12	3222.4	97.0	4580	6	AK037303
13	2623	78.9	2994	6	AK037303
14	2623	78.9	2994	6	AK037303
15	2623	78.9	2994	6	AK037303
16	2623	78.9	2994	6	AK037303
17	2623	78.9	2994	6	AK037303
18	2618.4	78.8	2994	6	AK037303
19	2284.6	68.8	4593	6	EL2560
20	2283	68.7	4591	10	MUS3195RB
21	2273.4	68.4	4432	10	KATRP
22	1473.2	44.3	3456	5	000113
23	1465	44.1	4541	5	GDRE11NB
24	1394	42.0	1554	4	AF230742
25	1035.8	31.2	2808	5	NVPETINBP
26	835	25.1	2352	6	159392
27	835	25.1	2352	9	HUMHRAIRA
28	835	25.1	2352	9	AK037303
29	835	25.1	180388	9	HUMHRAIRA
30	723.4	21.8	1815	11	006460
31	711.8	21.4	3486	5	AF102861
32	639.2	19.2	2730	5	AY098289
33	483.2	14.5	2102	10	KATRTNBLST
34	347.2	10.4	426	9	AF043224
35	288.6	8.7	944	5	AY125075
36	248	7.5	301	9	F336015S25
37	248	7.5	301	11	G73432
38	225.2	6.8	110188	2	AF105877
39	220.8	6.6	901	6	E12562
40	220.8	6.6	1698	6	E12561
41	220.8	6.6	3499	6	E12564
42	220.8	6.6	3633	6	E12565
43	218.8	6.6	1697	6	E09941
44	207	6.2	480	9	HUMHRAIRA
45	199	6.0	320	9	F336015S16

ALIGNMENTS

Result 1	HUMHRAIRA	Human retinoblastoma associated (RB) mRNA, complete cds.	4600 bp	Linear	PRO 12 JUN 1995
HUMHRAIRA					
LOCUS					
DEFINITION					
ACCESSION	M3647.102994				
VERSION	M3647.1	GI190945			
KEYWORDS	retinoblastoma protein.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
REFERENCE					
AUTHORS					
TITLE					

tumors: organization of the sequence and its encoded protein
Proc. Natl. Acad. Sci. U.S.A. 84 (24), 9059-9063 (1987)

JOURNAL
88097427

MEDLINE

PubMed

3480540

Drift entry and computer-readable copy of sequence for [1] kindly

provided by S H Friend, 10-PBR-1988

location/qualifiers

1..4600

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="13q14.2"

/clone="p4.7R"

/cell-type="retinoblast"

1..4600

/gene="RB1"

4..2790

/gene="RB1"

/note="retinoblastoma associated protein"

/codon_start=1

/protein_id="AA06486.1"

/db_xref="GI:190946"

/db_xref="AF08-300-118-734"

/translation="MPKTPKTKATATAAAAAEPAPPEPPPEDEDSPELPL

VRPEFETEPERFALCKLIKIPDVRKFWALIMKSVSSVKEVILGYTKKKKELWGLP

IFIAVDLDEMPFIFLQKQKILFVRRKFNILKELHDSIAVDNMSRLKRLVDLEA

IFSKLERICELIVLQPSISSIFINSAIVKVSIFILLAKCEVILQWHDIVISPOL

MLCVLDVIFIKLSPMLLKPEPKYKTAIVPINKSPPTPRQRNSARTAKOLENDRIEIV

LCKEHNIDEVNVKFNIPKMSNLGIVTSNGIPVENISKPYEFLYLKMKDLQAP

LEHDHKLITLSDSEFQRTPRKSNLDEENVVIPPHTPVTVNTIQGLMILNSA

SDQSTNIVSTFNNIVNKFSTIKKVKDQSTIPKPKANAQDQCVFGSPKTKLVV

KLYRWMSIMLKSEPERLSIQNFSLKLDNDIIFHMSLLACALEVVMATYSRTSONLDS

GUDISPTIIVNINIKAFDHYKVFISFKARGNIPKMIKIKHFRFHRIMSLAWISD

SPFLDIQSKDPRGPTDLESAPLNIPLNNHTAADMVLSVPSRPKKSTTPVNS

TANAEATSAFQTKPLKSTLSLEYKKVYLURLNTLKERLSEPEDEHLIMT

LFQHTLQNEVELDRDLQIMCMSYGVCKVKNIDLFKLIIVTAYKDLPRVAOETEK

KVLIKRPYSIIIVKYSVPMQIKTNIIQYASTPRPILSPHPIKPSYKPKSSPIK

IPGNIYISDIKSPYKISGPIPTPRKTPMPSRILVSIQHSFQTSSEKPKINOMVNSD

PVTKPSAPDSINPKPKIKKIFHIFQSDHAIQSKHIDPSKPKDKRMISTIPMKNG

KMDSMDINSKEEK"

BASE COUNT 1489 a 842 c 815 q 1454 t

REGION

Query Match 99.88; Score 3315; DB 9; Length 4600;

Rest Local Similarity 100 DB; Word No 6;

Matches 3315; conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 GGAAAGAAATTCATGCTGATGAGTATTTGGAGAGTTATATTCACAAAGAAAGAAACT 68

DB 237 GGAGAAAGATTCATGCTGATGAGTATTTGGAGAGTTATATTCACAAAGAAAGAAACT 296

QY 69 GTGGAAATCTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 128

DB 297 GTGGAAATCTATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 356

QY 129 TCAGTATACAGAAACATACAGATGATGATGATGATGATGATGATGATGATGATGATG 188

DB 457 TCAGTATACAGAAACATACAGATGATGATGATGATGATGATGATGATGATGATGATG 416

QY 189 TCATATACATACAGAAATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 248

DB 317 TCATATACATACAGAAATTCATGATGATGATGATGATGATGATGATGATGATGATG 476

QY 249 GTTGGATATCTTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCA 308

DB 477 GTTGGATATCTTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAG 536

QY 309 CAGTGGATATCTTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCA 368

DB 537 CAGTGGATATCTTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAG 596

QY 469 TTTATATACATACAGCAAAATTCATACAGCAAAATTCATACAGCAAAATTCATACAG 428

DB 597 TTTATATACATACAGCAAAATTCATACAGCAAAATTCATACAGCAAAATTCATACAG 656

QY 429 AATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488

DB 657 AATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716

QY 489 ATATAAAC 548

DB 717 ATATAAAC 776

QY 549 CAGAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTAC 608

DB 777 CAGAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGT 836

QY 609 TAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668

DB 837 TAAAGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 896

QY 669 TTTATGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 728

DB 897 TTTATGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 956

QY 729 TAAAGAT 788

DB 957 TAAAGAT 1016

QY 789 TAT 848

DB 1017 TAT 1076

QY 849 AAAAGAT 908

DB 1077 AAAAGAT 1136

QY 909 TATGAAAT 968

DB 1137 TATGAAAT 1196

QY 969 AAAAT 1028

DB 1137 AAAAT 1256

QY 1029 ACTGAAAGAT 1088

DB 1257 ACTGAAAGAT 1316

QY 1089 TGTGAAAT 1148

DB 1317 TGTGAAAT 1376

QY 1149 ATGAT 1208

DB 1377 ATGAT 1436

QY 1209 TCAGCAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1268

DB 1437 TCAGCAAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1496

QY 1269 TACAGAGAT 1328

DB 1497 TACAGAGAT 1556

QY 1329 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1388

DB 1557 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1616

QY 1389 AGAAGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1448

DB 1617 AGAAGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1676

QY 1449 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1508

DB 1677 GAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1736

BASE COUNT 1544 a 902 c 880 g 1523 t
ORIGIN
Query Match 99.88; Score 3315; DB 6; Length 4839;
Res: Local Similarity 100.0%, Pred. No. 0,
Matches 3315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 CGATAAAGTTCATCTGCTGGATGAGTATGGGAGGCTATATCAAAAGAAAAAGCAACT 68
DB 471 CGACAAGATTCATCTGCTGGATGAGTATGGGAGGCTATATCAAAAGAAAAAGCAACT 430
QY 69 GTGGAGAAATCTATCTTTATTTAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 128
DB 431 GTGGAGAAATCTATCTTTATTTAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT 490
QY 129 TCAATACAGAAAAACATACAAATACAGTTCCTTAAATATCTTTTAACTTAAACAAAT 188
DB 491 TCAATACAGAAAAACATACAAATACAGTTCCTTAAATATCTTTTAACTTAAACAAAT 550
QY 189 TCAATACAGAAAAACATACAAATACAGTTCCTTAAATATCTTTTAACTTAAACAAAT 248
DB 551 TCAATACAGAAAAACATACAAATACAGTTCCTTAAATATCTTTTAACTTAAACAAAT 610
QY 249 GTTTGCAATCTTCAAGAAATTTGAGAGGATGTGAGACATATATATTTGACACAAACAG 308
DB 611 GTTTGCAATCTTCAAGAAATTTGAGAGGATGTGAGACATATATATTTGACACAAACAG 670
QY 309 TGGTGGCAATCTTCAAGAAATTTGAGAGGATGTGAGACATATATATTTGACACAAAC 368
DB 671 TGGTGGCAATCTTCAAGAAATTTGAGAGGATGTGAGACATATATATTTGACACAAAC 730
QY 369 TTTTATACCTAAAGGGAAGTATTAATAAATGAGATGATGCTGCTGATTCATTCACAT 428
DB 731 TTTTATACCTAAAGGGAAGTATTAATAAATGAGATGATGCTGCTGATTCATTCACAT 790
QY 429 AATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488
DB 791 AATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850
QY 489 ATATAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 548
DB 851 ATATAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 910
QY 549 CAGGAGTGCAGGATACGAAACAGCTGAGAAATGATGATGATGATGATGATGATGATGAT 608
DB 911 CAGGAGTGCAGGATACGAAACAGCTGAGAAATGATGATGATGATGATGATGATGATGAT 970
QY 609 TAAATACATGAATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 668
DB 971 TAAATACATGAATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030
QY 669 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
DB 1031 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1090
QY 729 TAAATACATGAATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 788
DB 1091 TAAATACATGAATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1150
QY 789 TCAATACAGAAAAACATACAAATACAGTTCCTTAAATATCTTTTAACTTAAACAAAT 848
DB 1151 TCAATACAGAAAAACATACAAATACAGTTCCTTAAATATCTTTTAACTTAAACAAAT 910
QY 849 AAAAGTAAAGTTCATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 908
DB 1211 AAAAGTAAAGTTCATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1270
QY 909 TACAAATACATGAATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 968
DB 1271 TACAAATACATGAATGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1330
QY 969 AAAATGATTTCTATTTTAACTTAAACAAATTTTAACTTAAACAAATTTTAACTTAAACAA 1028

2411 GGGTCTTCAAGACAGACAGCAAAA'AAA'ATATTTGGAGATAGTTCACACAGGAGGAGCTAC 2470
 2419 CTGTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2168
 2471 CTGTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2530
 2499 GATTTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2328
 2531 GATTTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2530
 2529 TTTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2288
 2591 TTTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2650
 2489 ATTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2348
 2651 ATTGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2710
 2449 GTTAAAAAATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2408
 2711 GTTAAAAAATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2770
 2409 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2468
 2771 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2830
 2469 GTTAAAAAATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2528
 2831 GTTAAAAAATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2890
 2529 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2588
 2891 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2950
 2589 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2648
 2951 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3010
 2649 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2708
 3011 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3070
 2709 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2768
 3071 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3130
 2769 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2828
 3131 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3190
 2829 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2888
 3191 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3250
 2889 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 2948
 3251 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3310
 2949 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3008
 3311 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3370
 3009 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3068
 3371 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3430
 3069 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3128
 3431 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3490
 3129 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3188
 3491 TATGTCATATATCTCAATATCTGAGAGGCTTAT'AAATTTCTATCTATCTAGCTTACG 3550

3189 TAAAGCTGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4248
 3551 TAAAGCTGGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4610
 3249 ACCGACAGATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4308
 3611 ACCGACAGATGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4670
 3309 AGGATATTGATAGT 4323
 3671 AGGATATTGATAGT 4685
 RESULT 3
 HUMBLIMRNA
 LOCUS
 DEFINITION Homo sapiens retinoblastoma susceptibility protein (Rb) mRNA and mutations.
 ACCESSION L41870.1 GI:794994
 VERSION 1
 KEYWORDS retinoblastoma protein; retinoblastoma susceptibility.
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 McGeer, J.L., Yandell, D.W. and Dryja, T.P.
 Structure and partial genomic sequence of the human retinoblastoma susceptibility gene
 Gene 80 (1), 119-128 (1989)
 90086771
 2701949
 2 (sites)
 Rbq.A., Guadim, Z., Baird, P.N. and Cowell, J.K.
 Detection of heterozygous mutations in the Rb1 gene in retinoblastoma patients using single strand conformation polymorphism analysis and polymerase chain reaction sequencing
 Oncogene 7 (7), 1445-1451 (1992)
 92419557
 1352398
 3 (sites)
 Guadim, Z., Baird, P.N. and Cowell, J.K.
 Oncogenic point mutations in exon 20 of the Rb1 gene in families showing incomplete penetrance and mild expression of the retinoblastoma phenotype
 Proc. Natl. Acad. Sci. U.S.A. 89 (14), 6177-6181 (1992)
 92435261
 1352883
 4 (sites)
 Guadim, Z., Baird, P.N. and Cowell, J.K.
 Mechanisms of oncogenesis in patients with familial retinoblastoma with bilateral and unilateral retinoblastoma
 Proc. Natl. Acad. Sci. U.S.A. 90 (15), 7451-7455 (1993)
 93348271
 8217609
 5 (sites)
 Rbq.A., Bia, R., Guadim, Z. and Cowell, J.K.
 Molecular mechanisms of oncogenic mutations in tumors from patients with bilateral and unilateral retinoblastoma
 Proc. Natl. Acad. Sci. U.S.A. 90 (15), 7451-7455 (1993)
 93348271
 8346255
 6 (sites)
 Kralzke, R.A., Peterson, G.A., Boock, A., Coxon, A.R., Geradts, J., Cowell, J.K. and Kaye, F.J.
 Partial inactivation of the Rb product in a family with incomplete penetrance of familial retinoblastoma and benign retinal tumors
 Oncogene 9 (5), 1321-1326 (1994)
 94203660
 8152792
 Mutation data provided by Dr. B.L.Gallie, Hospital for Sick children 555 Univ. Ave.


```

QY 2824 AAAATGCTGCTGTTTATGCAATAGCAATGCGGCTGACGTGCGATGCGATGCAATGCAATG 2888
DB 2824 AAAATGCTGCTGTTTATGCAATAGCAATGCGGCTGACGTGCGATGCGATGCAATGCAATG 3113
QY 2869 GCGTGGTGGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2948
DB 3114 GCGTGGTGGTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3173
QY 3244 TTAATTTATGATGATATTTTAAATTAACATGACACACCTTAGAAAATGCTGCTATC 3008
DB 3244 TTAATTTATGATGATATTTTAAATTAACATGACACACCTTAGAAAATGCTGCTATC 3233
QY 3309 TATCTGCGCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3068
DB 3244 TATCTGCGCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3293
QY 3369 CAAAAAGCAATATATGACAAATGACAAAAATGATGATGATGATGATGATGATGATGATGAT 3128
DB 3244 CAAAAAGCAATATATGACAAATGACAAAAATGATGATGATGATGATGATGATGATGATGAT 3353
QY 3429 TTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3188
DB 3429 TTTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3413
QY 3489 TAAAAATGGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3248
DB 3414 TAAAAATGGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3473
QY 3549 ACCGACAAATGCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3308
DB 3414 ACCGACAAATGCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3533
QY 3609 AGGATTTATGATGAT 3423
DB 3544 AGGATTTATGATGAT 3548

```

RESULT 1
 Length 109469
 DEFINITION Sequence 1 from Patent WO 8906703.
 ACCESSION 109469
 VERSION 109469.1 GI:587899
 KEYWORDS
 SOURCE unknown.
 ORGANISM unknown.
 REFERENCE 1 (bases 1 to 4597)
 AUTHORS Dryja J.P., Friend S. and Yandell D.W.
 TITLE DIAGNOSIS OF RETINOBLASTOMA
 JOURNAL Patent. WO 8906703 A 1 27 JUL 1989.
 FEATURES
 Location/Qualifiers
 1..4597
 /organism="unknown"
 BASE COUNT 1489 a 841 c 813 g 1454 t
 BEGIN

Query Match 99.7% Score 3313.4 DB 6 Length 4597;
 Best Local Similarity 100.0% Pred. No. 0;
 Matches 4314. Conservative 0. Mismatches 1. Indels 0. Gaps 0;

```

QY 9 GCAAAAATTTTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 58
DB 24 GCAAAAATTTTCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 293
QY 49 GCGGGAATGCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 128
DB 24 GCGGGAATGCTGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 353
QY 129 TGAGTATGAAAAACATGAAAAACATGAAAAACATGAAAAACATGAAAAACATGAAAAACAT 188
DB 44 TGAGTATGAAAAACATGAAAAACATGAAAAACATGAAAAACATGAAAAACATGAAAAACAT 413

```

```

QY 189 TCATACGACATGACAAATTCATATGCTATGCAAGACATGTCAGAAAGTATGATGATGAT 248
DB 414 TGATACGACATGACAAATTCATATGCTATGCAAGACATGTCAGAAAGTATGATGATGAT 473
QY 219 GTTTCGACATGTCAGCAAAATTCGCAAAAGCAATGCTGCAACTATATATATTCACACAAATG 308
DB 474 GTTTCGACATGTCAGCAAAATTCGCAAAAGCAATGCTGCAACTATATATATTCACACAAATG 533
QY 309 CAGTTCATATCTACTCTAATAAATTCCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 458
DB 534 CAGTTCATATCTACTCTAATAAATTCCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 593
QY 369 TTTATTAGCTAAAGCTGCAACTATTAATAAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 428
DB 594 TTTATTAGCTAAAGCTGCAACTATTAATAAGCAATGATGCTGCTGCTGCTGCTGCTGCTGCTG 553
QY 429 AATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 488
DB 654 AATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
QY 489 ATATAAAGCACTGTTATAGCAATTAAGTATGCACTGCAAGCAATGCTGCTGCTGCTGCTGCT 548
DB 714 ATATAAAGCACTGTTATAGCAATTAAGTATGCACTGCAAGCAATGCTGCTGCTGCTGCTG 773
QY 549 CAGGAGTCGACGCGATACCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 608
DB 774 CAGGAGTCGACGCGATACCAAAAGCAATGATGATGATGATGATGATGATGATGATGATGATG 833
QY 609 TAAACAACATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 668
DB 834 TAAACAACATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 893
QY 669 TTTTATCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
DB 894 TTTTATCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 953
QY 729 TAAAGCATGACGAAATTTATCTTAAAAATTAAGATGATGATGATGATGATGATGATGATGAT 788
DB 954 TAAAGCATGACGAAATTTATCTTAAAAATTAAGATGATGATGATGATGATGATGATGATGAT 1013
QY 789 TCATCATGAAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 848
DB 1014 TCATCATGAAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
QY 849 AAAAAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 908
DB 1074 AAAAAGTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1133
QY 909 TATGACACATCTGCAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 968
DB 1134 TATGACACATCTGCAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1193
QY 969 AAATCTGATCTGCTGATTTTAAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1028
DB 1194 AAATCTGATCTGCTGATTTTAAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253
QY 1079 AGTCAGGATATAGGATATATCTTTTAAAGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1088
DB 1254 AGTCAGGATATAGGATATATCTTTTAAAGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
QY 1089 GCTGCAAAATGGAATGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1148
DB 1314 GCTGCAAAATGGAATGCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1373
QY 1149 ATCCATGCTTAAATGCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
DB 1374 ATCCATGCTTAAATGCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1433
QY 1209 TGACAAATTTTTCATATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1268
DB 1434 TGACAAATTTTTCATATGCTTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1493

```


DB 187 TGTGAACTTATATATTTTCACACAAACAGGAGTTCGATATCTTACTCAAAATAAATCTTGCA 246
QY 438 TTGGTGTAAAGATTTTGTGATACATTTTATTATGCTAAAGAGGAGTATTACAAATG 399
DB 247 TTGGTGTAAAGATTTTGTGATACATTTTATTATGCTAAAGAGGAGTATTACAAATG 306
QY 400 CAAGATGATCTGGGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 459
DB 407 GAAGATGATCTGGGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 366
QY 410 CTTCTACATCTGGGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 519
DB 467 CTTCTACATCTGGGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 426
QY 520 TCACTCTGAACATCTGAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 579
DB 427 TCACTCTGAACATCTGAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 486
QY 580 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 639
DB 487 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 546
QY 430 AAAAAATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 699
DB 547 AAAAAATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 606
QY 700 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 759
DB 407 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 666
QY 760 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 819
DB 467 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 726
QY 820 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 879
DB 727 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 786
QY 880 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 939
DB 987 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 846
QY 940 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 999
DB 947 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 906
QY 1000 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1059
DB 907 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 966
QY 1060 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1119
DB 967 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1026
QY 1120 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1179
DB 1027 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1086
QY 1180 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1239
DB 1087 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1146
QY 1240 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1299
DB 1147 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1206
QY 1300 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1359
DB 1207 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1266
QY 1360 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1419
DB 1419 AATGATATTAAGATTTTATTTTACGTTAAATGCTATGCTGCTTCTGCTACTATTTTATTA 1419

DB 1267 TACAAAGTATCGAAAGCTTTTATCAAAAGCAGAGGCAACTTCAGAACACACAAATCAAAAA 1326
QY 1420 CATTTATTAATATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1479
DB 1327 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1386
QY 1480 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1539
DB 1387 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1446
QY 1540 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1599
DB 1447 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1506
QY 1600 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1659
DB 1507 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1566
QY 1660 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1719
DB 1567 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1626
QY 1720 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1779
DB 1627 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1686
QY 1780 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1839
DB 1687 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1746
QY 1840 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1899
DB 1747 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1806
QY 1900 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1959
DB 1807 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1866
QY 1960 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2019
DB 1867 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1926
QY 2020 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2079
DB 1927 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 1986
QY 2080 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2139
DB 1987 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2046
QY 2140 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2199
DB 2047 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2106
QY 2200 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2259
DB 2107 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2166
QY 2260 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2319
DB 2167 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2226
QY 2320 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2379
DB 2227 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2286
QY 2380 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2439
DB 2287 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2346
QY 2440 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2499
DB 2347 CATTTATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAATATTAAT 2406

16 2999 TTTGGTAAAAATTTCTTGGATACATTTTTATTAGTAAAGGGAGATTACAAATG 2931
17 400 GAAGAAGATCGGGAATTCATTCACATTAAGCTTAAGTCTCTGACATTTATTAATA 459
18 2940 GAAGATGATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2871
19 460 CTCTACATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 519
20 2870 CTCTACATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2811
21 520 TAACTGGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 579
22 2810 TAACTGGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2751
23 580 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 639
24 2750 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2691
25 640 AAAAATGTTTATTTAAAAATTTTATAGCTTTTATGAAATCTCTTGGACTGTGAAATCT 699
26 2690 AAAAATGTTTATTTAAAAATTTTATAGCTTTTATGAAATCTCTTGGACTGTGAAATCT 2631
27 700 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 759
28 2640 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2571
29 760 AAAAATGTTTATTTAAAAATTTTATAGCTTTTATGAAATCTCTTGGACTGTGAAATCT 819
30 2570 AAAAATGTTTATTTAAAAATTTTATAGCTTTTATGAAATCTCTTGGACTGTGAAATCT 2511
31 820 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 879
32 2510 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2451
33 880 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 939
34 2450 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2391
35 940 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 999
36 2390 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2331
37 1000 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1059
38 2330 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2271
39 1060 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1119
40 2270 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2211
41 1120 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1179
42 2210 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2151
43 1180 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1239
44 2150 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2091
45 1240 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1299
46 2090 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 2031
47 1300 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1359
48 2030 GTGAATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1971
49 1360 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1419
50 1970 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1911
51 1420 AATGATATCGGGAATTTCTTCAATTCATTCAGTTAAGCTATGCTGCTGCTATTTATTA 1479

Db 1910 CATTAGAGGATGTGAATATGGAATGCTTGAATGATGATGATGATGATGATGATGATGAT 1861
Qy 1480 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1539
Db 1850 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1791
Qy 1540 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1599
Db 1790 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1741
Qy 1600 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1659
Db 1740 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1691
Qy 1660 GAT 1719
Db 1670 GAT 1611
Qy 1720 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779
Db 1610 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
Qy 1780 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1839
Db 1550 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491
Qy 1840 GAT 1899
Db 1490 GAT 1441
Qy 1900 TATGAT 1959
Db 1440 TATGAT 1471
Qy 1960 AATGAT 2019
Db 1370 AATGAT 1411
Qy 2020 TATGAT 2079
Db 1310 TATGAT 1291
Qy 2080 TATGAT 2139
Db 1250 TATGAT 1191
Qy 2140 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2199
Db 1190 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1141
Qy 2200 GAT 2259
Db 1130 GAT 1071
Qy 2260 TATGAT 2319
Db 1070 TATGAT 1011
Qy 2320 AATGAT 2379
Db 1010 AATGAT 951
Qy 2380 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2439
Db 950 CTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
Qy 2440 AATGAT 2499
Db 890 AATGAT 831
Qy 2500 AATGAT 2559
Db 830 AATGAT 771

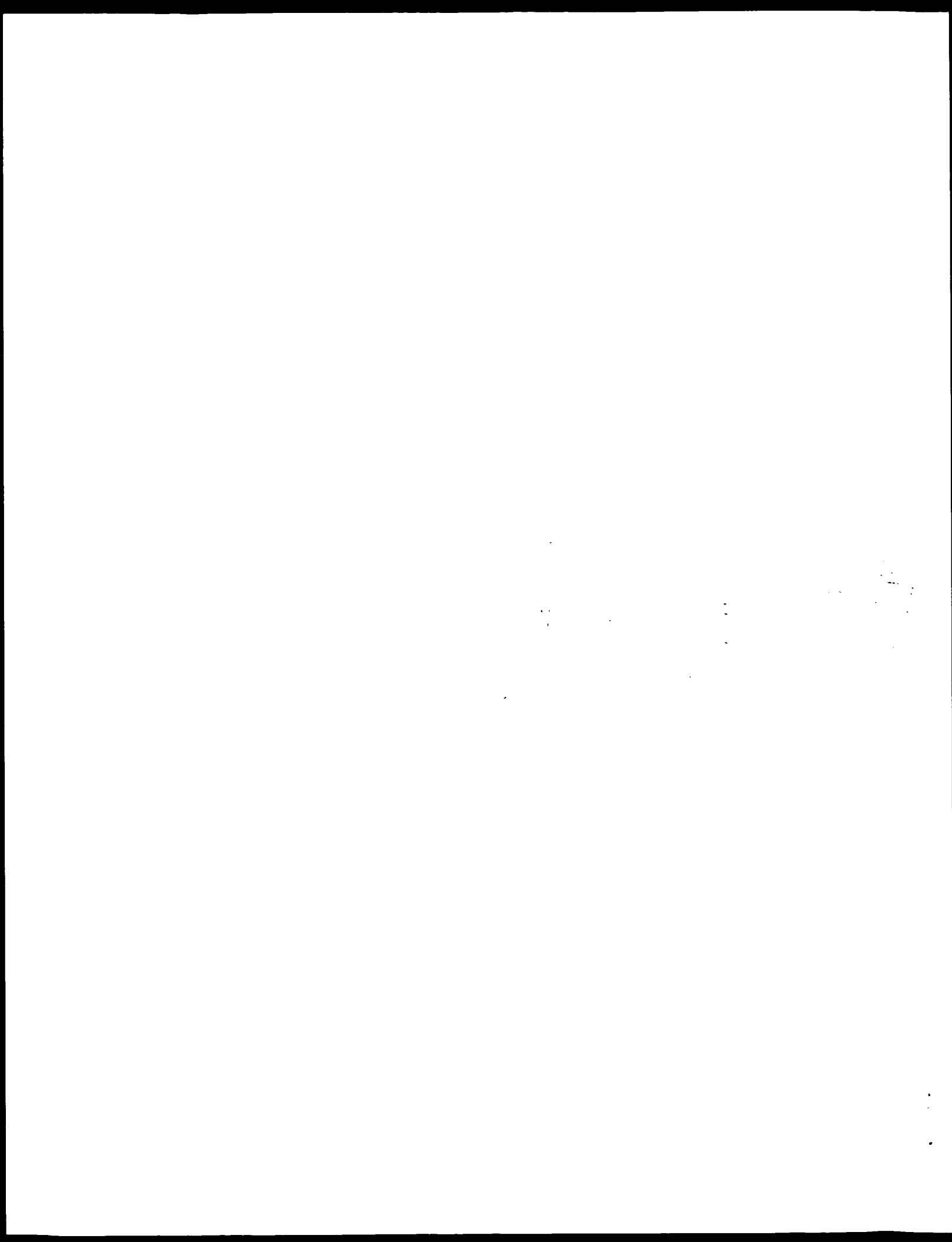
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

1632 TAGCAAGTAATCTCGAATCTTATCTTCTGAAAGATTTGTTTCTTCTGCAUGATTCT 1691
1329 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1388
1692 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1751
1389 AGAAGGCAATTTGCAAGAGAAATGATAAACATTTAGAACGATGTAAATATGAATATAT 1448
1752 AGAAGGCAATTTGCAAGAGAAATGATAAACATTTAGAACGATGTAAATATGAATATAT 1811
1449 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1508
1812 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1871
1509 AGAAGGCAATTTGCAAGAGAAATGATAAACATTTAGAACGATGTAAATATGAATATAT 1568
1872 AGAAGGCAATTTGCAAGAGAAATGATAAACATTTAGAACGATGTAAATATGAATATAT 1931
1569 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1628
1932 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1991
1629 AGAAGGCAATTTGCAAGAGAAATGATAAACATTTAGAACGATGTAAATATGAATATAT 1688
1992 AGAAGGCAATTTGCAAGAGAAATGATAAACATTTAGAACGATGTAAATATGAATATAT 2051
1689 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1748
2052 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2111
1749 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1808
2112 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2171
1809 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1868
2172 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2231
1869 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1928
2232 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2291
1929 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 1988
2292 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2451
1989 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2048
2352 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2411
2049 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2108
2412 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2471
2109 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2168
2472 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2531
2169 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2228
2532 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2591
2229 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2288
2592 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2651
2289 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2348
2652 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2711
2349 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2408
2712 GAATGCTCTTAATTTAAAGCCCTTTGATTTTACAAAGTCATCCAAAGTTTATCAAAAT 2771

QY 949 CAGCAGTGGAGGATAGCAAAATACAAAAAGATACAGCAATTAATGAGTCTCTG 608
DB 912 CAGCAGTGGAGGATAGCAAAATACAAAAAGATACAGCAATTAATGAGTCTCTG 971
QY 609 TAAACAAATACAGCAATTAATGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 668
DB 912 TAAACAAATACAGCAATTAATGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1031
QY 669 TTTATGAAATCT 728
DB 1032 TTTATGAAATCT 1091
QY 729 TAAATGAAATCT 788
DB 1092 TAAATGAAATCT 1151
QY 789 TAAATGAAATCT 848
DB 1152 TAAATGAAATCT 1211
QY 849 TAAATGAAATCT 908
DB 1212 TAAATGAAATCT 1271
QY 909 TAAATGAAATCT 968
DB 1272 TAAATGAAATCT 1331
QY 969 TAAATGAAATCT 1028
DB 1332 TAAATGAAATCT 1391
QY 1029 TAAATGAAATCT 1088
DB 1392 TAAATGAAATCT 1451
QY 1089 TAAATGAAATCT 1148
DB 1452 TAAATGAAATCT 1511
QY 1149 TAAATGAAATCT 1208
DB 1512 TAAATGAAATCT 1571
QY 1209 TAAATGAAATCT 1268
DB 1572 TAAATGAAATCT 1631
QY 1269 TAAATGAAATCT 1328
DB 1632 TAAATGAAATCT 1691
QY 1329 TAAATGAAATCT 1388
DB 1692 TAAATGAAATCT 1751
QY 1389 TAAATGAAATCT 1448
DB 1752 TAAATGAAATCT 1811
QY 1449 TAAATGAAATCT 1508
DB 1812 TAAATGAAATCT 1871
QY 1509 TAAATGAAATCT 1568
DB 1872 TAAATGAAATCT 1931
QY 1569 TAAATGAAATCT 1628
DB 1932 TAAATGAAATCT 1991

QY 1629 AACTGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 1688
DB 1992 AACTGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2051
QY 1689 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 1748
DB 2052 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2111
QY 1749 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 1808
DB 2112 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2171
QY 1809 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 1868
DB 2172 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2231
QY 1869 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 1928
DB 2232 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2291
QY 1929 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 1988
DB 2292 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2351
QY 1989 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2048
DB 2352 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2411
QY 2049 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2108
DB 2412 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2471
QY 2109 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2168
DB 2472 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2531
QY 2169 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2228
DB 2532 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2591
QY 2229 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2288
DB 2592 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2651
QY 2289 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2348
DB 2652 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2711
QY 2349 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2408
DB 2712 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2771
QY 2409 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2468
DB 2772 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2831
QY 2469 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2528
DB 2832 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2588
QY 2529 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2588
DB 2892 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2951
QY 2589 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2641
DB 2952 CACAGAGCGCTGAAATTTACTGCAAAATGACACACACACACACACACACACACAC 2994

Search completed: January 17, 2003, 18:06:18
Job time : 5546.29 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

QM protein nucleic search, using frame_plus_p2n model

Run on: January 17, 2003, 18:15:29 : Search time 408.78 Seconds
(without alignments)
6678.212 Million cell updates/sec

Files: us-09-026-459a-31
Perfect score: 4504
Sequence: 1 MDTALQKLKIPUHVPERA.....TRMKKMKNSMUTSNKEEK 874

Scoring table:

BLAST/BL62
Xgapop 10.0 : Xgapext 0.5
Ygapop 10.0 : Ygapext 0.5
Fgapop 6.0 : Fgapext 7.0
Bgapop 6.0 : Bgapext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4104280

Minimum DB seq length: 0

Maximum DB seq length: 60000000

Post processing: Minimum Match 100%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

Model: frame_plus_p2n model -DEXT-0

Lib: GenEmbl -QMT-tastap -SUFFIX-rpe -MINMATCH-0.1 -LOCAL-0 -LOCALEXT-0

Units-bits -START-1 -END-1 -MATRIX-BLOSUM62 -TRANS-human40.cdi -LIST-45

LOCALIGN-200 -THR SCORE-PCT -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL

OUTFILE-pro -NORM-ext -HAPSIZ-500 -MINLEN-0 -MAXLEN-200000000

DBP: US09020459a-31-324000000 -152137-23309 -NC100-6 -IC92-7

NO XLEXY NO HMAP -LAP-QUERY -NEG_SCORES-0 -WALT -CONJ-02 -DEXT-120

WARN TIMEOUT 40 -THREHDS-1 -XGAPOP-10 -XGAPEXT-0.5 -F5GAPEXT-7

YGAPOP-10 YGAPEXT-0.5 DELTAP-6 DELEXT-7

Database:

- 1: db_ba:*
- 2: db_btq:*
- 3: db_btq:*
- 4: db_btq:*
- 5: db_btq:*
- 6: db_btq:*
- 7: db_btq:*
- 8: db_btq:*
- 9: db_btq:*
- 10: db_btq:*
- 11: db_btq:*
- 12: db_btq:*
- 13: db_btq:*
- 14: db_btq:*
- 15: db_btq:*
- 16: db_btq:*
- 17: db_btq:*
- 18: db_btq:*
- 19: db_btq:*
- 20: db_btq:*
- 21: db_btq:*
- 22: db_btq:*
- 23: db_btq:*
- 24: db_btq:*
- 25: db_btq:*
- 26: db_btq:*
- 27: db_btq:*
- 28: db_btq:*

- 29: em_vit:*
- 30: em_btq_hum:*
- 31: em_btq_inq:*
- 32: em_btq_other:*
- 33: em_btq_mus:*
- 34: em_btq_pln:*
- 35: em_btq_rnd:*
- 36: em_btq_mam:*
- 37: em_btq_vit:*
- 38: em_sy:*
- 39: em_btqo_hum:*
- 40: em_btqo_mus:*
- 41: em_btqo_other:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4499	99.9	2994	6	AKG91965 Sequence
2	4499	99.9	2994	6	AKG9189 Sequence
3	4499	99.9	2994	6	AKG9189 Sequence
4	4499	99.9	2994	6	AKG9189 Sequence
5	4499	99.9	2994	6	AKG9189 Sequence
6	4499	99.9	2994	6	AKG9189 Sequence
7	4499	99.9	2994	6	AKG9189 Sequence
8	4499	99.9	2994	6	AKG9189 Sequence
9	4499	99.9	2994	6	AKG9189 Sequence
10	4499	99.9	2994	6	AKG9189 Sequence
11	4499	99.9	2994	6	AKG9189 Sequence
12	4499	99.9	2994	6	AKG9189 Sequence
13	4499	99.9	2994	6	AKG9189 Sequence
14	4499	99.9	2994	6	AKG9189 Sequence
15	4499	99.9	2994	6	AKG9189 Sequence
16	4499	99.9	2994	6	AKG9189 Sequence
17	4499	99.9	2994	6	AKG9189 Sequence
18	4499	99.9	2994	6	AKG9189 Sequence
19	4499	99.9	2994	6	AKG9189 Sequence
20	4499	99.9	2994	6	AKG9189 Sequence
21	4499	99.9	2994	6	AKG9189 Sequence
22	4499	99.9	2994	6	AKG9189 Sequence
23	4499	99.9	2994	6	AKG9189 Sequence
24	4499	99.9	2994	6	AKG9189 Sequence
25	4499	99.9	2994	6	AKG9189 Sequence
26	4499	99.9	2994	6	AKG9189 Sequence
27	4499	99.9	2994	6	AKG9189 Sequence
28	4499	99.9	2994	6	AKG9189 Sequence
29	4499	99.9	2994	6	AKG9189 Sequence
30	4499	99.9	2994	6	AKG9189 Sequence
31	4499	99.9	2994	6	AKG9189 Sequence
32	4499	99.9	2994	6	AKG9189 Sequence
33	4499	99.9	2994	6	AKG9189 Sequence
34	4499	99.9	2994	6	AKG9189 Sequence
35	4499	99.9	2994	6	AKG9189 Sequence
36	4499	99.9	2994	6	AKG9189 Sequence
37	4499	99.9	2994	6	AKG9189 Sequence
38	4499	99.9	2994	6	AKG9189 Sequence
39	4499	99.9	2994	6	AKG9189 Sequence
40	4499	99.9	2994	6	AKG9189 Sequence
41	4499	99.9	2994	6	AKG9189 Sequence
42	4499	99.9	2994	6	AKG9189 Sequence
43	4499	99.9	2994	6	AKG9189 Sequence
44	4499	99.9	2994	6	AKG9189 Sequence
45	4499	99.9	2994	6	AKG9189 Sequence

ALIGNMENTS

RESULT 1


```

1144 1144 TTTTGGATCATGATAAAATCTTTCACATCATCTTATACACAGCTTTTCAACACACAGACA 1203
QY 402 ThrProArgLysSerAsnLeuAspGluGluValAsnValIleProProHisThrProVal 321
Db 1204 ACACACGCAAAAAGTAACCTTGTATGAAGAGGTGAATGTAAATCTCTCCACACACTCCAGTT 1263
QY 422 ArgThrValMetAsnThrIleGlnGlnIleuMetMetIleuAsnSerAlaSerAspGln 341
Db 1264 AGGACTATTATGAATATATGAAATATATATATATATATATATATATATATATATATATAT 1323
QY 442 ProSerGluAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProLysGluSerIle 361
Db 1324 GCTTTCAGAAAAATCTGATTTTCATTTTAAACATTCGACATGATTCCTTCAAGAGTATA 1383
QY 462 LeuGlnAsnValLysAsnIleGlyTyrIlePheLysLysPheAlaLysAlaValGly 381
Db 1384 CTGAAGAGTGAAGATATAGGATACATCTTTAAACAGAAATTTGCTAAAGCTGAGGGA 1443
QY 482 GlyGlyLysValIleLysSerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArg 401
Db 1444 CAGAGTTGTGTGAAATTTGATCACAATATATATATATATATATATATATATATATATATAT 1503
QY 492 ValMetGlnSerMetLeuLysSerGluGluGluArgGluSerIleGlnAsnProSerLys 421
Db 1504 GAAATGCAATTCATGCTTAAATCAGAAAGAACAGATTTCCATTCACAAATTTTAGCAAA 1563
QY 422 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMet 441
Db 1564 CTTCTGCAATCAATCAATTTTCAATGCTTTATGCGCGCGCTGCGAGGTGTAATG 1623
QY 442 AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro 461
Db 1624 GCAATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1683
QY 462 ThrIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGluSerPhe 481
Db 1684 TCGATCTGAAATGCTTAAATTTAAAGCCCTTGTATTTTACAAAGTGAATGCAAGATTTT 1743
QY 482 ThrLysAlaThrLysLeuThrArgGluMetIleLysHisLeuGluArgCysGluHis 501
Db 1744 ATCAAAATATCAAGTAATTTTCAAAATGAAATGATATATATATATATATATATATATATAT 1803
QY 502 ArgIleMetGluSerLeuAlaThrLeuSerAspSerProLeuPheAspLeuIleLysGln 521
Db 1804 GGAATCATGGAATCTGCTTATGATATGATATGATATGATATGATATGATATGATATGATAT 1863
QY 522 SerLysAspArgGluGlyProThrAspHisLeuSerAlaCysProLeuAsnLeuPro 541
Db 1864 TCAAGGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1923
QY 542 LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 561
Db 1924 CTCCAGATATATCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1983
QY 562 LysGlnSerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 581
Db 1984 AAGATTTTAAATATATATATATATATATATATATATATATATATATATATATATATATATAT 2043
QY 582 LeuGlnHisThrLysSerAlaLysSerAlaLysSerAlaLysSerAlaLysSerAla 601
Db 2044 TCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2103
QY 602 ArgLeuAlaTyrLeuArgLeuAsnThrLeuGlyGluArgLeuSerGluHisProGlu 621
Db 2104 GAGATAGATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2163
QY 622 LeuGlnHisThrLysSerAlaLysSerAlaLysSerAlaLysSerAlaLysSerAla 641
Db 2164 TTAATAATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2223
QY 642 ArgAspArgHisLeuAspGlnIleMetMetCysSerMetTyrCylLysLysValLys 661

```

```

Db 2224 ACACACGCAATTTTGGACCAAAATATATCATCTCTTCCATCTATGCGCAATCCAAATTCAG 2283
QY 562 AsnIleAspLeuLysPheLeuLysIleLeuValThrAlaIleLysAspLeuProGluAlaVal 581
Db 2284 AATATATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2343
QY 682 GlnGluThrPheLysArgValLeuIleLysGlnGlnSerIleLysAspSerIleIleValPhe 701
Db 2344 CAGAGACATTCATATATATATATATATATATATATATATATATATATATATATATATATAT 2403
QY 702 TyrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArg 721
Db 2404 TATAATCTGCTTCTTCAACACACATCAAAACAAATATATTCACATAGCTTCTTCTATATAT 2463
QY 722 ProProThrLeuSerProIleProHisIleProArgSerProTyrLysPheProSerSer 741
Db 2464 GAGCTATCTTCTTCAACACACATCAAAACAAATATATTCACATAGCTTCTTCTATATAT 2523
QY 742 ProLeuArgIleProGlyLysIleTyrIleSerProLeuLysSerProTyrLysIle 761
Db 2524 CATTATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2583
QY 762 SerGluGlyLeuProThrArgThrLysMetThrProArgSerArgIleLeuValSerIleP 781
Db 2584 TCAGAGCTCTGCAACACACCAACAAATTCACATCAAGATCAAGAAATCTAGTATATATAT 2643
QY 782 GlyGluSerThrGlyThrSerGluLysPheGluLysLeuGluMetValGluAsnSer 801
Db 2644 GGTGATATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2703
QY 802 AspArgValLysLysArgIleGlySerPheSerPheSerPheSerPheSerPheSerPheSer 821
Db 2704 GAGGTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2763
QY 822 ArgPheAspIleGlySerAspGluAlaAspGluSerLysLeuArgGlySerLysSer 841
Db 2764 CATTATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2823
QY 842 LysPheGlnGlnLysLeuAlaGluMetThrSerThrArgMetGlnLysGlnLys 861
Db 2824 AATTTTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCAAAATTCAGCA 2883
QY 862 MetAsnAspSerMetAspThrSerAsnLysGluGluLys 874
Db 2884 ATCAATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2922

```

RESULT 3
AR207831 Sequence 3 from patent US 6379927.

LOCUS AR207831 2994 bp DNA 11.0.2002
DEFINITION AR207831
ACCESSION AR207831
VERSION AR207831.1 GI:21507687
KEYWORDS
SOURCE Unknown.
ORIGIN
REFERENCE 1 (bases 1 to 2994)
AUTHORS Antelmaz, D., Gregory, P.J., and Willis, K.N.
TITLE Retinoblastoma fusion proteins
JOURNAL Patent, US 6379927 A 30 Apr 2002
FEATURES
SOURCE 1: 2994
/organism="unknown"

BASE COUNT 9/4 a 618 c 593 g 809 t
ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2994
Score: 4499.00 Matches: 873
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.89% Indels: 0
DB: 6 Gaps: 0


```

QY 722 ProThrLeuSerProIleProHisIleProArgSerProTyrLysPheProSerSer 741
DB 2464 CCCCCIACCTIGICACCAATACCTACATCCCTCGAACCCCTTACAGTTCTCTACTTCA 2523
QY 742 ProLeuArgIleProGlyGlyAsnIleTyrLysSerProLeuLysSerProTyrLysIle 761
DB 2524 CCGTTACGATCTCTGAGGAAATATATATATTTATATATATTTATATATATTTATAT 2583
QY 762 SerGluGlyLeuProThrProThrLysMetThrProArgSerArgIleLeuValSerIle 781
DB 2584 TCAGAGAGCTGCGCAACACCAACCAAAATGACTCCAGATCAAGATCTTAGTATCAATT 2643
QY 782 GlyIleSerPheGlyThrSerGlyLysPheGlnLysIleAsnGlnMetValCysAsnSer 801
DB 2644 GCGCAATCATGCGGACATCTGCAAGATCTCCACAAATAAAATACATCTATCTATCTAAC 2703
QY 802 AspArgValLeuLysArgSerAlaGlyGlySerAsnProProLysProLeuLysLysLeu 821
DB 2704 GATGCTCTCTTAAATTAATATGATGATGATGATGATGATGATGATGATGATGATGAT 2763
QY 822 ArgPheAspIleGluGlySerAspGluAlaAspGlySerLysHisLeuProGlyGluSer 841
DB 2764 GCTTTGATATGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2823
QY 842 LysPheGlnGlnLysLeuAlaGlnMetThrSerThrArgMetThrArgMetGlnLysGlnLys 861
DB 2824 AAAATTCACCAACATGCGACAGAAATGACTTCTACCGAACACGAAATGCCAAAGCAGAAA 2883
QY 862 MetAsnAspSerMetAspThrSerAspLysGluGluLys 874
DB 2884 ATGAATGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2922

RESULTS 4
R0009728
L0005
DEFINITION Tissue specific expression of 1-41 in fibrosarcoma protein
ACCESSION B0009728
VERSION B0009728.1 GI:18638101
FEATURES
SOURCE JP 2001504638-A/2.
ORGANISM unidentifed.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Antomovic, Gregory, P. I. and Willis, K. N.
TITLE Tissue specific expression of retinoblastoma protein
JOURNAL CANJ J INC
COMMENT
OS unidentifed
PN JP 2001504638-A/2
PD 21-MAR-2001
PR 13-NOV-1997 JP 1998522958
PR 15-NOV-1996 JP 092751517, 14-DEC-1997 US 09/280199, 01
DEPOSITED IN ANTIMAN, PICHARD, J. GREGORY, KENNETH N. WILLS, PG
CC(=) 3.04. CC(=) 3.04. CC(=) 3.04. CC(=) 3.04. CC(=) 3.04. CC(=) 3.04.
CC Strandedness: Single;
CC Topology: Linear;
FH KEY location/Qualifiers
FI source 1..2994
FT Location/Qualifiers
1..2994
/oranism="unidentifed"
/DB_xref="taxon:32644"
BASE 1 (IN) 474 a 618 c 593 g 809 t
- R0009728

Alignment Scores:
Prod. No.: 0 Length: 2994
Score: 4499.00 Matches: 873
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.89% Indels: 0

```

```

DB: 6 Gaps: 0
US-09-026-459a-31 (1-874) x H0009728 (1-2994)
QY 2 AspPheThrAlaLeuCysClnLysLysLysIleProAspHisValArgGluArgAlaIleP 21
DB 304 GATTTTACTGCTATATGTCAGTAATTAAGATATATATATATATATATATATATATATAT 363
QY 22 LeuPheThrProGluLysValSerValAspGlyValLeuGlyGlyTyrIleClnLysLys 41
DB 364 TTAACCTTGGACAGAAAGTTTCACTGCGATGGAGTATTCGGAGCTTATATTCAAACAAA 123
QY 42 LysGlnLeuIleThrGlyIleGlyValPheIleAlaAlaValAspLeuAspGlnMetSerPhe 61
DB 424 AAGCAACGTGGGCAATCTGATCTTTATTCGACGCACTTCACCTAGATGAGATGCGTTC 483
QY 62 ThrPheThrGluLeuGlnLysAsnIleGluIleSerValHisLysPhePheAsnLeuLeu 81
DB 484 ACITTTACTGAGCTACAGAAAAACATAGAAATCAGTGTGCTAATAATCTTTAACTTACTA 543
QY 82 LysGluIleAspThrSerThrLysValAspAspAlaMetSerAlaLeuLeuLysLysLys 101
DB 544 AAAGAAATTCATACCCAGTACCAAGTTGATATGCTATGCTCAAGACTGTTCAGAGAT 603
QY 102 AspValLeuPheAlaLeuPheSerLysLeuGluArgThrCysGluLeuIleTyrLeuThr 121
DB 604 CATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 663
QY 122 ClnProSerSerSerIleSerThrGlnIleAsnSerAlaLeuValLeuLysValSerIleP 141
DB 664 CAGCCAGAGCTGGATATCTGTAATAATAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 723
QY 142 IleThrPheLeuLeuAlaLysGlyGluValLeuGlnMetClnAspAspLeuValIleSer 161
DB 724 ATCACTTTTATAGCTAAAGGGAGATATACAAATCGAACAACATGCTGCTGCTGCTGCT 783
QY 162 PheGlnLeuMetLeuLysValLeuAspTyrPheIleLysLeuSerProMetLeuLeu 181
DB 784 TTTCAGTTAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
QY 182 LysGlnProTyrLysThrAlaValIleProPheIleAsnGlySerProArgThrProArg 201
DB 844 AAGAACCATATAAACAGACGTTATACCCATTATGTTGCTGCTGCTGCTGCTGCTGCTGCT 903
QY 202 GlyIleAsnArgSerAlaArgIleAlaLysClnLysGlnAsnAspThrArgIleCln 221
DB 904 GGTTCAGAACACAGAGTCCACATAGCAAAACACATAGAAATCATACAGAATATTATCAA 963
QY 222 ValIleGlyLysGlnIleThrLysAsnIleAspGlnValLysAsnValTyrPheLysAsn 241
DB 964 GTTCTGCTGTAACCAACATCAATGTAATATAGATGAGGTTGAAAAATGTTTATTCAAAA 1023
QY 242 PheIleThrPheLeuMetAspSerLeuGlyLeuValThrSerAspGlyLeuThrGluVal 261
DB 1024 TTTAATATTTATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1083
QY 262 AsnLeuSerLysArgTyrGluGluIleTyrLeuLysAsnLysAspLeuAspAlaArgLeu 281
DB 1084 AATCTTTCTAAACGATACGAAGAAATTTATCTTTAAAAATAAAGATCTACATGCAACA 1143
QY 282 PheLeuAspHisAspLysThrLeuGlnIleThrAspSerIleAspSerPheLeuThrGln 401
DB 1144 TTTTGTGATCATGATAAAACCTCTTCAGACTGATGCTATACAGACTTTTGAACACAGACA 1203
QY 302 ThrProArgLysSerAsnLeuAspGlyGluValAsnValIleProGlnHisThrProVal 421
DB 1204 ACACCAAGAAAAAGTAACCTTGAATGAAGAGTGAATGTAATGCTGCTGCTGCTGCTGCT 1263
QY 322 ArgThrValMetAsnThrIleGluGlnLeuMetMetIleLeuAsnSerAlaSerAspGln 341
DB 1264 AGGACTGTTAAGAACATATCCCAACAAATTAATCATGATTTAAATTCACCAAGATCAAA 1323
QY 342 ProSerGluAsnLeuIleSerTyrPheAspAsnGlyThrValAsnProLysGluSerIle 361

```

```

1424 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1483
1425
1426 LeuLeuSerValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 381
1427
1428 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1443
1429
1430 LeuLeuSerValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 401
1431
1432 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1503
1433
1434 ValMetGluSerMetLeuLeuSerGluGluGluGluGluGluGluGluGluGluGluGluGlu 421
1435
1436 GTAAATGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1563
1437
1438 LeuLeuAsnAspAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 441
1439
1440 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1623
1441
1442 AlaThrTyrSerArgSerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPhePro 461
1443
1444 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1683
1445
1446 TrpLeuLeuAsnValLeuAsnValLeuLysAlaPheAspPheTyrLysValIleGluSerPhe 481
1447
1448 TGAATTCGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATG 1743
1449
1450 LeuLysAlaGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 501
1451
1452 ATCAAGAGAGAAAGGAAATTTTACAAAGAAATGATTAATTTAGAACCATGTGACAT 1803
1453
1454 ArgAlaMetGluSerLeuAlaTrpLeuSerAspSerProLeuPheAspLeuLeuLysGln 521
1455
1456 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1863
1457
1458 SerLysAspArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 541
1459
1460 TCAAGAGAGAGAAAGGAAATTTTACAAAGAAATGATTAATTTAGAACCATGTGACAT 1923
1461
1462 LeuGluAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 561
1463
1464 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 1983
1465
1466 LysGluSerThrThrArgValAsnSerThrAlaAsnAlaGluThrGluAlaThrSerAla 581
1467
1468 AAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATG 2043
1469
1470 PheGluThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 601
1471
1472 TTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 2103
1473
1474 ArgAlaAlaValLeuArgAlaAsnThrLeuGluGluGluGluGluGluGluGluGluGluGlu 621
1475
1476 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 2163
1477
1478 LeuGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 641
1479
1480 TCAAGAGAGAGAAAGGAAATTTTACAAAGAAATGATTAATTTAGAACCATGTGACAT 2223
1481
1482 ArgAspArgHisLeuAspAlaIleMetMetCysSerMetTyrGlyIleCysLysValLys 661
1483
1484 AAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATG 2283
1485
1486 AsnLeuAspLeuLysPheLysIleValThrAlaTyrLysAspLeuProHisAlaVal 681
1487
1488 AAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATGCTGAAATG 2343
1489
1490 GluGluThrPheLysArgValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 701
1491
1492 CTTTCAGAAAATCTGATTTCCTATTTTAAACATGTCACAGTGAATCCAAAAAGAAATATA 2403
1493
1494 TyrAsnSerValPheMetGluArgLeuLysThrAsnIleLeuGluTyrAlaSerThrArg 721
1495

```

```

Db 2404 TATAATCTGGTCTTCATGTCAGAGATGAGAAAATAAATAATTTTGGAGATATGCTTCAATGAGG 2463
QY 722 ProProThrLeuSerProIleProHisIleProArgSerProIleLysPheProSer 741
Db 2464 CCGCTACCTTCTGACCAATATATCACAATCTGCGAAAGGCTTAAATGCTTCAATGAGG 2523
QY 742 ProLeuArgIleProProLysIleProLysIleProLysIleProLysIleProLysIle 761
Db 2524 CCGCTACCGAATCTGCGAGAGAAATATATATATATATATATATATATATATATATATAT 2583
QY 762 SerGluGluLeuProIleProIleThrLysMetIleProArgSerArgIleLeuValSerIle 781
Db 2584 TCGAAGGCTGCTGCAACATCAACAAAATATGATCGAATATCAATATATATATATATATATAT 2643
QY 782 GlyGluSerPheGlyThrSerGluLysPheGluLysIleAsnGluMetValCysAsnSer 801
Db 2644 GCGCAATATCTGCGACATCTGCAAGGCTGAGAAAATAAATATATATATATATATATATATAT 2703
QY 802 AspArgValLeuLysArgSerAlaGluGlySerAspIleProLysIleProLysIleProLys 821
Db 2704 CACTGCTGCTGCAAAAGAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCT 2763
QY 822 ArgPheAspIleGluGlySerAspGluAlaAspGlySerLysHisLeuProIleGluGluSer 841
Db 2764 CAGTTTGATATATGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2823
QY 842 LysPheGluGlnIleLysLeuAlaGluMetThrSerThrArgThrArgMetGluGluLysGln 861
Db 2824 AAATTCAGCAACAACTGCGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2883
QY 862 MetAsnAspSerMetAspThrSerAsnLysLysLysLysLysLysLysLysLysLysLys 874
Db 2884 ATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2922
RESIDUE 5
LOCUS 181465 2994 bp LINA Linear PAT 10-JUN 1998
DEFINITION Sequence 2 from patent: US 5712555.
ACCESSION 181465
VERSION 181465.1 GI:3204762
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Shepard,H,Michael, and Wen,S,Fen.
TITLE Characterization of a novel anti-p110.sup.KH monoclonal antibody
JOURNAL FASEB J 10:555-55 A 20 JAN 1998;
FEATURES
Source 1: 2994
/organism="unknown"
BASE COUNT 974 a 618 c 593 g 809 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 2994
Score: 4499.00 Matches: 874
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 99.89% Indels: 0
DB: 6 Gaps: 0
US-09-026-459A-31 (1-874) x 181465 (1-2994)
QY 2 AspheThrAlaLeuCysGlnIleLysLysIleProAspHisValArgGluAlaValPhe 21
Db 304 GATTTTACATGATATATGTCAGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 363
QY 22 LeuThrTrpGluLysValSerSerValAspGlyValLeuGlyGlyTrpIleGluGlyLys 41
Db 364 TTAACCTGGGAGAAAGTTTCACTGTCGCAATGCGAGATATATATATATATATATATATATATAT 423
QY 42 LysGluLeuTrpIleGlyIleGlyIlePheIleAlaAlaValAspLeuAspGluMetSerPhe 61

```

100 424 AAGCAATGCGGGAAATGCTGATGCTATATGACCACTTACCTGACATGAGATGCGCTC 483
42 ThrPheThrGluLeuGlnLysAsnIleGluSerValHisLysPhePheAsnLeuLeu 81
100 441 ACCTTACAGACCTACACAAACATACAAATCAGTGTCCATAAAATCTTTAACTTACTA 543
92 LysGluIleAspThrSerThrLysValAspAsnAlaMetSerArgGluCysLysLysVal 101
100 544 AAAAATATGATACAT 603
102 AspValIlePheAlaLeuPheSerLysLysGluArgThrCysGlnLeuIleLysLeuThr 121
100 604 GATGATATGCTTGGACATCTTCAGCAAAATGGAAGACACATGCAACTATATATATGACA 663
102 GlnProSerSerSerIleSerThrGluIleAsnSerAlaLeuValLeuLysValSerTyr 141
100 644 CAATCAGACATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATG 723
102 ThrPheLeuLeuAlaLysGlyGluValIleGlnMetGluAspAsnValIleSer 161
100 744 ATCAAT 783
102 PheGlnLeuMetLeuCysValLeuAspIlePheIleLysLeuSerProProMetLeuLeu 181
100 784 ATTCAT 843
102 LysLeuIleProLysThrAlaValIleProIleAspGlySerProArgThrProArgArg 201
100 844 AAAAATAATAAAAAATGCTGATATATATATATATATATATATATATATATATAT 903
202 GlyCysAspArgSerAlaArgIleAlaLysGlnLeuGluAsnAspThrArgIleGlu 221
100 904 GCTCAGAAATGAT 963
222 ValIleCysLysGlnHisGlyCysAsnIleAspGluValLysAsnValIleThrLysAsn 241
100 964 GTCTCTGTAAGAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1023
242 PheIleProPheMetAsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGlu 261
100 1024 TTTAT 1083
262 AsnIleCysLysArgThrGlyGlyIleThrLysLysAsnLysAspLeuAspArgLeu 281
100 1084 AATCTTTTAAACGATACGAATATATATATATATATATATATATATATATATATAT 1143
282 PheLeuAspHisAspLysThrLeuGlnThrAspSerIleAspSerPheGluThrGlnArg 301
100 1144 TTTTGGATCATGATAAAACGCTTCACATGATTCATATAGACACTTTTGAACACAGAGA 1203
302 ThrProArgLysSerAsnLeuAspGluGluValAsnValIleProProHisThrProVal 321
100 1204 ACAGCAAGAAACATACCTTGAAGAGAGGTGCAATGATATATATATATATATATAT 1263
322 ArgThrValMetAsnThrIleGlnIleGlnMetMetIleLeuAsnSerAlaSerAspGln 341
100 1264 AAT 1323
342 ProSerGlnAsnIleCysThrPheAsnAsnCysThrValAsnProLysGluSerIle 361
100 1324 GCTTCAAAAAATGAT 1383
362 LeuLysArgValLysAspIleGlyIlePhePheLysGlyLysPheAlaLysAlaValGly 381
100 1384 GCGAAACAGTCAAGGATATAGGATACATCTTTAAAGACAAATTTGCTAAAGCTGCGCA 1443
382 GlnGlyLysValGluIleGlySerGluArgGlyLysLeuGlyValArgLeuThrArg 401
100 1444 CAGATTTGTCAT 1503
402 ValMetGluSerMetLeuLysSerGluGluArgLeuSerIleGlnAsnPheSerLys 421

1504 GTATGCAATCCATGCTTAAATCAGAAAGAAAGCAATACCATTCAAAAATTTATACAAA 1563
422 LeuLeuAspAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMet 441
1564 GTTCAT 1623
442 AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerCysLysThrAspLeuSerPhe 461
1624 GGCACATATAGCAAGAAATATATATATATATATATATATATATATATATATATAT 1683
462 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheCysLysValIleGluSerPhe 481
1684 TGCATTCGAATGCTTAAATTTAAAGAGCTTGCATTTTACAAAGCTACCAAACTTTT 1743
482 IleLysAlaGluGlyAsnLeuThrArgGluMetIleLysHisLeuGluArgCysGluHis 501
1744 ATCAAAAT 1803
502 ArgIleMetGlnSerLeuAlaThrLeuSerAspSerProLeuPheAspLeuIleLysGln 521
1804 GCAAT 1863
522 SerLysAspArgGluGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuPro 541
1864 TCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1923
542 LeuGlnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 561
1924 CTCAGANTAT 1983
562 LysGlySerThrThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 581
1984 AAGGTTTAAAT 2043
582 PheLeuThrGlnLysProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyr 601
2044 TCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2103
602 ArgLeuAlaTyrLeuArgLeuAsnThrLeuCysGluArgLeuLeuSerGlnHisProGlu 621
2104 GGTAT 2163
622 LeuGluHisIleIleThrPhePheGlnHisThrLeuGluAsnGlyTyrGluLeuMet 641
2164 TTAAT 2223
642 ArgAspArgHisLeuAspGlnIleMetMetCysSerMetTyrGlyIleCysLysValLys 661
2224 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2283
662 AsnIleAspLeuLysPheLysIleValThrAlaTyrLysAspLeuProHisAlaVal 681
2284 AT 2343
682 GlnGluThrPheLysArgValLeuIleLysGluGlyCysThrAspSerIleIleValPhe 701
2344 CAGCAAT 2403
702 TyrAsnSerValPheMetGlnArgLeuLysThrAsnIleLeuGlyTyrAlaSerThrArg 721
2404 TATAACGCTGCTTCAAGCAGACACTCAAAACAAATATTTTCCAGATGCTTCAAGTGT 2463
722 ProProThrLeuSerProIleProHisIleProArgSerProLysPheProSerSer 741
2464 CAGAT 2523
742 ProLeuArgIleIleProGlyAsnIleTyrIleSerProLeuLysSerProLysLys 761
2524 CATTAT 2583
762 SerGluLysLeuProThrProThrLysMetThrProArgSerArgIleLeuValSerIle 781
2584 TCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2643


```

QY 662 ThrProAlaGlySerAsnLeuAspGluValAsnValIleProProHisThrProVal 321
DB 1204 ATATAGAGCTTTAAATTCAAAATCATTTGTAACAGCATATAGAGATCTGCTCATGTGTT 2443
QY 682 GlnGluThrPheLysArgValLeuIleLysGluGluGluTyrAspSerIleIleValPhe 701
DB 2344 CAGCAGACATTCAAACCTGTTTTCATCAACAAACAGCAGATGATGATATATATATATAT 2403
QY 702 TyrAsnSerValPheMetGlnLeuLysThrAsnIleLeuGlnTyrAlaSerThrArg 721
DB 2404 TATAACTGATCTTATATGACAGACATCAAAATTAATATTTTCATGATATGATATATATAT 2463
QY 722 ProProThrLeuSerProIlePheHisIleProArgSerProTyrLysPheProSerSer 741
DB 2464 GATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2523
QY 742 ProLeuArgIleProGlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIle 761
DB 2524 CCTTAGCGATTCCTGAGGAGCAATCTATATATTCACCTGAGGAGCAATCTATATATATAT 2583
QY 762 SerGluGlyLeuProThrProThrLysMetThrProArgSerArgIleLeuValSerIle 781
DB 2584 TAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2643
QY 782 GlyGluSerPheGlyThrSerGluLysPheGlyLysIleAsnGluMetValLysAsnSer 801
DB 2644 GATCAATATTGAGGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2703
QY 802 AspArgValLeuLysArgSerArgGluGlySerAsnProProLysProLeuLysLysLeu 821
DB 2704 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2763
QY 822 ArgPheAspIleGluLysSerAspGluAlaAspGlySerGlyHisIleLeuProGlyLys 841
DB 2764 GATTTTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2823
QY 842 LysPheGlnGlnLysLeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLys 861
DB 2824 AATTTTCAGCAGAAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2883
QY 862 MetAsnAspSerMetAspThrSerAsnLysGluLys 874
DB 2884 ATCAATCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2922

RESULT 7
105311
LOCUS SCIENCE 1 from Patent EP 0259031.
DEFINITION 105311
ACCESSION 105311
VERSION 105311.1 G1:591083
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 4597)
TITLE Dryja, J.P. and Friend, S.
TITLE Human DNA in the diagnosis of retinoblastoma
FEATURES
FEATURE 1 (bases 1 to 4597)
LOCATION/Qualifiers
SOURCE 1. 4597
BASE COUNT 1489 a 842 c 812 g 1454 t
ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 4597
Score: 4499.00 Matches: 873
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.84% Indels: 0
DB: 6 Gaps: 0
US-09-026-459A-31 (1-874) x 105311 (1-4597)

```

```

QY 662 ThrProAlaGlySerAsnLeuAspGluValAsnValIleProProHisThrProVal 321
DB 1204 ATATAGAGCTTTAAATTCAAAATCATTTGTAACAGCATATAGAGATCTGCTCATGTGTT 2443
QY 682 GlnGluThrPheLysArgValLeuIleLysGluGluGluTyrAspSerIleIleValPhe 701
DB 2344 CAGCAGACATTCAAACCTGTTTTCATCAACAAACAGCAGATGATGATATATATATATAT 2403
QY 702 TyrAsnSerValPheMetGlnLeuLysThrAsnIleLeuGlnTyrAlaSerThrArg 721
DB 2404 TATAACTGATCTTATATGACAGACATCAAAATTAATATTTTCATGATATGATATATATAT 2463
QY 722 ProProThrLeuSerProIlePheHisIleProArgSerProTyrLysPheProSerSer 741
DB 2464 GATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2523
QY 742 ProLeuArgIleProGlyGlyAsnIleTyrIleSerProLeuLysSerProTyrLysIle 761
DB 2524 CCTTAGCGATTCCTGAGGAGCAATCTATATATTCACCTGAGGAGCAATCTATATATATAT 2583
QY 762 SerGluGlyLeuProThrProThrLysMetThrProArgSerArgIleLeuValSerIle 781
DB 2584 TAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2643
QY 782 GlyGluSerPheGlyThrSerGluLysPheGlyLysIleAsnGluMetValLysAsnSer 801
DB 2644 GATCAATATTGAGGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2703
QY 802 AspArgValLeuLysArgSerArgGluGlySerAsnProProLysProLeuLysLysLeu 821
DB 2704 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2763
QY 822 ArgPheAspIleGluLysSerAspGluAlaAspGlySerGlyHisIleLeuProGlyLys 841
DB 2764 GATTTTATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2823
QY 842 LysPheGlnGlnLysLeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLys 861
DB 2824 AATTTTCAGCAGAAATGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2883
QY 862 MetAsnAspSerMetAspThrSerAsnLysGluLys 874
DB 2884 ATCAATCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2922

```



```

variation
179 /gene="RB1"
    /note="Tumor GOS564 unilateral mutation causing a tentative
    wrong splice; G00-118-734"
    /citation-[5]
    /replace-"t"
    275...401
    /gene="RB1"
    /note="G00-118-734"
    /number-2
    402...517
    /gene="RB1"
    /note="G00-118-734"
    /number-3
    506...509
    /gene="RB1"
    /note="RNF64 bilateral familial deletion mutation causing
    premature stop"
    /replace-"ca"
    518...637
    /gene="RB1"
    /note="G00-118-734"
    /number-4
    537...542
    /gene="RB1"
    /note="Tumor GOS561 unilateral frameshift mutation causing
    premature stop"
    /citation-[5]
    /replace-"ta"
    621...622
    /gene="RB1"
    /note="Tumor GOS537 bilateral frameshift mutation causing
    premature stop; G00-118-734"
    /citation-[5]
    /replace-"c"
    638...676
    /gene="RB1"
    /note="G00-118-734"
    /number-5
    677...744
    /gene="RB1"
    /note="G00-118-734"
    /number-6
    745...855
    /gene="RB1"
    /note="G00-118-734"
    /number-7
    856...998
    /gene="RB1"
    /note="G00-118-734"
    /number-8
    944...1076
    /gene="RB1"
    /note="G00-118-734"
    /number-9
    1077...1186
    /gene="RB1"
    /note="G00-118-734"
    /number-10
    1095
    /gene="RB1"
    /note="Tumor GOS559 mutation causing premature stop"
    G00-118-734"
    /citation-[5]
    /replace-"t"
    1104
    /gene="RB1"
    /note="Tumor GOS563 unilateral mutation causing premature
    stop; G00-118-734"
    /citation-[5]
    /replace-"t"
    1187...1264
    /gene="RB1"

```

```

    /note="G00-118-734"
    /number-11
    1269
    /gene="RB1"
    /note="Tumor GOS159 unilateral mutation causing a
    premature stop; G00-118-734"
    /citation-[5]
    /replace-"t"
    1265...1352
    /gene="RB1"
    /note="G00-118-734"
    /number-12
    1320
    /gene="RB1"
    /note="Tumor GOS554 mutation causing premature stop;"
    G00-118-734"
    /citation-[5]
    /replace-"t"
    1352...1375
    /gene="RB1"
    /note="Tumor GOS13 unilateral frameshift mutation
    resulting in premature stop. Associated with this mutation
    is a large 13q14.3 sub-band deletion; G00-118-734"
    /citation-[5]
    /replace-"ca"
    1353...1469
    /gene="RB1"
    /note="G00-118-734"
    /number-13
    1400...1401
    variation

```

Alignment Scores:

```

Pred. No.: 0 length: 4839
Score: 4499.00 Matches: 874
Percent Similarity: 100.00% Conservatize: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.89% Indels: 0
DB: 9 Gaps: 0

```

US-09-026-459a-31 (1-974) x HUMRB1.MRNA (1-4839)

```

Q7 2 AsppPheThrAlaLeuCysGlcLysLeuLysLysIleProAspPheValAlaArgAlaArgAlaTrp 21
Db 303 GATTTTATGTAATATGTCAGAAATTAAGATACCAATCATGTACAGAGAGAGGCTTGG 462
Q7 22 TGGPThrGluLysValSerValAspGlyValIleuGlyGlyValIleGlyLysLys 41
Db 363 TTAACTTGGGAGAAATTTTCATCTGTGTGATGAGATATGAGAGAGGTTATATTTAAAAGAAA 422
Q7 42 LysGluLeuTrpGlyIleGlySerPheIleValAlaValAspLeuAspGluMetSerPhe 61
Db 423 AAGCAACTCTGGTAAATCTGATCTTTATTCACCACTTCACCTACAGACATGCTGCTTC 492
Q7 62 ThrPheThrGluLeuGlnLysAsnIleGluIleSerValIleLysPhePheAsnLeuLeu 81
Db 483 ACTTTTATTAAGATATAGAAAAATATAGAAALAGTGGCATTAATATCTTTTACCTTATTA 542
Q7 82 LysGluLeuAspThrSerThrLysValAspAsnAlaMetSerArgGluGluLysLysLys 101
Db 543 AAAGAAATTCATATCAGTACCAAGTTGATATGCTATGTCACCACTTCACCACTACAGAGAT 402
Q7 102 AspValLeuPheAlaLeuPheSerLysLeuGluArgThrCysGluLeuIleTrpLeuThr 121
Db 603 GAGGTATTTCTTGTATCTTCACCAAAATGGAAAGACATCTCAACTATATATTTACATA 642
Q7 122 GlnProSerSerSerIleSerThrGluIleAspSerAlaLeuValLeuLysValSerTrp 141
Db 664 CAATATTAAGATATATATATATATATATATATATATATATATATATATATATATATAT 722
Q7 142 ThrPheLeuLeuAlaLysGlyGluValIleGluLeuMetGluAspAspGluValIleSer 161
Db 723 ATCAATTTTATATATATATATATATATATATATATATATATATATATATATATATATAT 782

```

```

162 PheGlnLeuMetLeuCysValIleuAspTyrPheTleCysLeuSerProPheMetLeuLeu 181
163 TTTCAATTAAATGATATGTTGCTTGAATATTATTAAACTCTCAGCTCCATGTTGCTC 842
164
165 LysGluProTyrTyrLeuAlaValIleProIleAsnGlySerProGlnIleProGlnArg 201
166 AAAGAAATATAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAATGATATG 902
167
168 GlyIleAsnArgSerAlaAlaIleAlaTyrGlnLeuIleAsnAspThrArgIleIleGlu 221
169 GGTATGAAATAGGAGTGTAAAGATACAAATACAAATACAAATACAAATACAAATAT 962
170
171 ValIleCysGlyGluHisGluCysAsnIleCysGluValLysAsnValTyrPheLysAsn 241
172 GTTATGTTGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAATGATAT 1022
173
174 PheIleProIleMetAsnSerLeuGlyIleValThrSerAsnGlyLeuIleGlnValGlu 261
175 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1082
176
177 ArgGlnSerLysArgTyrIleGlnIleTyrLeuLysAsnLysAspLeuAspAlaArgLeu 281
178 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1142
179
180 PheGlnAspHisAspLysThrLeuGlnIleThrAspSerIleAspSerPheGlnIleArg 301
181 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1202
182
183 ThrProArgLysSerAsnLeuAspGlnIleValAsnValIleProProIleThrProVal 321
184 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1262
185
186 ArgThrValMetAsnThrIleGlnIleGlnIleMetMetIleLeuAsnSerAlaSerAspGln 341
187 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1322
188
189 ProSerGlnAsnLeuIleSerLysPheAsnAsnCysThrValAsnProLysGluSerIle 361
190 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1382
191
192 LeuGlyArgValLysArgIleGlyTyrIlePheLysSerLysPheValLysAlaValGly 381
193 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1442
194
195 GlnGlyCysValGlnIleLysSerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArg 401
196 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1462
197
198 ValMetGluSerMetLeuLysSerGlnGlnIleValArgLeuSerIleGlnAsnPheSerLys 421
199 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1502
200
201 LeuLeuAsnAsnAsnIlePheHisMetSerLeuLeuAlaLysAlaGlnValValMet 441
202 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1562
203
204 AlaIleThrSerArgSerThrSerGlnAsnLeuLeuAspSerGlyThrAspLeuSerPhePro 461
205 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1622
206
207 TrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrLysValIleGlnAsnPhe 481
208 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1682
209
210 LysGlnAlaGlnIleGlnAsnLeuThrArgGlnMetIleLysHisLeuGlnArgCysGlnHis 501
211 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1742
212
213 ArgIleMetGluSerLeuAlaIleTrpLeuSerAspSerProLeuPheAspLeuIleLysGln 521
214 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1802
215
216 SerLysAspArgGlnGlyProIleProIleAspHisLeuGlnSerAlaCysProLeuAsnLeuPro 541

```

RES01.11
109469

```

1863 TCAAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1922
1864
1865 LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProLysLys 561
1866 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 1982
1867
1868 CysGlySerThrArgValAsnSerThrAlaAsnAlaGlnIleThrGlnAlaIleSerAla 581
1869 AAAGTTCATTAATGAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAGTGTAAAG 2042
1870
1871 PheGlnThrGlnLysProIleLysSerThrSerLeuSerLeuPheTyrTyrLysValTyr 601
1872 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2102
1873
1874 ArgLeuAlaTyrLeuArgAlaLeuAspThrLeuLysAlaAsnLeuLeuSerGlnIleLeuGln 621
1875 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2162
1876
1877 CysGlnArgCysGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 641
1878 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2222
1879
1880 ArgAspArgHisLeuAspGlnIleMetMetCysSerMetTyrGlyIleLysLysValLys 661
1881 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2282
1882
1883 AsnIleAspLeuLysPheLysIleIleValThrAlaTyrLysAsnLeuProIleHisAlaVal 681
1884 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2342
1885
1886 GlnGlnThrPheLysArgValIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 701
1887 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2402
1888
1889 TyrAsnSerValPheMetLeuArgLeuLysThrAsnIleLeuGlnIleValAlaSerThrArg 721
1890 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2462
1891
1892 ProProThrLeuSerProIleProIleHisIleProArgSerProTyrLysPheProSerSer 741
1893 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2522
1894
1895 ProLeuArgIleProGlnLysAsnIleLysIleSerProLeuLysSerProTyrLysIle 761
1896 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2582
1897
1898 SerGlnGlyLeuProIleProIleLysMetThrProArgSerArgIleLeuValSerIle 781
1899 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2642
1900
1901 GlyGlnSerPheGlnThrSerGlnLysPheGlnIleGlnIleGlnIleGlnIleGlnIle 801
1902 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2702
1903
1904 AspArgValLeuLysArgSerAlaGlnLysSerAsnThrProLysProLeuLysLysLeu 821
1905 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2762
1906
1907 ArgPheAspIleGlnIleLysSerAspGlnAlaAspGlySerLysHisLeuProGlnSer 841
1908 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2822
1909
1910 LysPheGlnIleLysLeuAlaGlnMetThrSerThrArgThrAspMetGlnGlyAlaLys 861
1911 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2882
1912
1913 MetAsnAspSerMetAspIleThrSerAsnLysGlnLysLys 874
1914 TTTTAAATGATATGTTAAAGAAATATGTTAAATGATATGTTAAATGATATGTTAAAT 2942
1915

```

```

100469 100469 4597 bp DNA Linear PAT 02-DEC-1994
DEFINITION Sequence 1 from Patent WO 8906703.
ACCESSION 100469
VERSION 100469.1 GI:587800
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 4597)
Dryja T.P., Friend S. and Yandell D.W.
TITLE DIAGNOSIS OF RETINOBLASTOMA
JOURNAL Patent: WO 8906703-A 1 27-Jul-1989;
FEATURES
Location/Qualifiers
1..4597
/orqanism="unknown"
BASE COUNT 1489 a 841 c 813 g 1454 t
ORIGIN
1
2 AspPheThrAlaLeuGlyGlnLysLeuLysPheProAspHisValArgGluArgAlaTrp 21
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2603
2604
2605
2606
2607
2608
2609
2610
2611
2612
2613
2614
2615
2616

```


1144 TTTTGGATCATGATATAAAGATCTTACACTGATTTATAGACAGATTTTGAACACACAGAGA 1203

QY 602 ThrProAlaGlySerAsnLeuAspGluValAsnValIleProHisThrProVal 321

1204 ACAATATGAAAGTAACCTTGAAGAAGCTGAATGTAATCTCCACACACATCCAGTT 1263

QY 622 ArgThrValMetAsnThrIleGluGluLeuMetIleLeuAsnSerAlaSerAspGln 341

1204 AGGACGCTTATCAACACATATCCACAAATTAATGATGATTTTAAATCCACCAAGATCAA 1323

QY 642 ProSerGlnAsnLeuIleSerTyrPheAsnAsnCysThrValAsnProIysGluSerIle 361

1444 GCTTCAGAAAATCTGATTTGCTATTTTAAAGAACTGCACAGTGAATCCAAAGAAATATA 1383

QY 662 LeuIysAspIleValIysAspIleCysThrIleHisLeuGlySerPheAlaValGly 381

1484 CCGAAAACAGTCAAGCAATATAGCAATATCTTTAAAGAGAAATTCGTAAGCGTGGCA 1443

QY 682 GlyGlyValGlyIleGlySerGlnArgTyrIleLeuGlyValArgLeuTyrArg 401

1444 CAGGCTGCTGCGAAATGGCATCACCGCATACAAACTTCGAGTTCGCTGTATACCCA 1503

QY 692 ValMetGluSerMetLeuIysSerGluGluGluArgLeuSerIleGlnAsnPheSerIys 421

1504 GTAATGGAATCCATGCTTAAATCAGAGAGAGAGGATTTCCCATTCAAAATTTTACAAA 1563

QY 692 LeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAlaCysAlaLeuGluValMet 441

1544 CTTTGTGAAATGACAACTTTTCAATATGCTTTATTCGGTGGCGGCTTCGAGGTGTAATG 1623

QY 692 AlaThrTyrSerArgSerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPhePro 461

1544 GGCATATATACAGCAAGTACATGCAAGATCTTCATTTCTGCAACAGATTTGCTTTCCCA 1683

QY 692 ThrIleLeuAsnValLeuAsnLeuLysAlaPheAspPheTyrIysValIleGluSerPhe 481

1604 TGGATTTGAAATGCTTAATTTAANAAGCTTTCATTTTACAAATGATGCGAAGTTT 1743

QY 692 IleIysAlaGluGlyAsnLeuThrArgGlnMetIleIysHisLeuGluArgCysGluHis 501

1644 ATCAAAATATAGAAATATATGATATATGATATATGATATATGATATATGATATATGAT 1803

QY 692 ArgIleMetGluSerLeuAlaTyrLeuSerAspSerProLeuPheAspLeuIleIysGln 521

1644 GCAATATATGAAATGCTTATGATATATGATATATGATATATGATATATGATATATGAT 1863

QY 692 SerIysAspArgGluGlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuPro 541

1664 TCAAAAGACCGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1923

QY 692 LeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSerProValArgSerProIysIys 561

1694 CACCAATATATGATATATGATATATGATATATGATATATGATATATGATATATGATAT 1983

QY 692 TysIysSerThrArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAla 581

1704 AAAGCTTCAACACCGCTGATAATTTCTACGCGAAATCCACACACACACACACACAC 2043

QY 692 PheGlnThrGlnIysProIysCysSerThrSerLeuSerLeuPheTyrIysValTyr 601

1744 TTGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2103

QY 692 ArgGlnAlaTyrLeuArgLeuAsnThrIleCysCysGluArgLeuSerGluHisProGlu 621

1704 GCGGTAAGCTATATGCGGCTAATATATATATATATATATATATATATATATATATAT 2163

QY 692 LeuGlnHisIleIleThrThrIlePheGluHisThrIleGlnAsnLeuTyrGlnLeuMet 641

1744 ATACAAATAAT 2223

QY 692 ArgAspArgHisIleuAspGlnIleMetMetCysSerMetTyrGlyIleCysIysValIys 661

1724 ATACAAATAAT 2283

QY 662 AsnIleAspLeuLysPheLysIleIleValIleAlaTyrIleLysAspLeuProHisAlaVal 681

Db 2284 AATATACACCTTAAATTCAAAATCAATGTAAACAGCAACAAGGATCTCTCAIGCTG 2343

QY 682 GlnGluThrPheLysArgValLeuIleLysGluGluGlyArgSerIleIleValPhe 701

Db 2344 CAGAGAAATTCAAAAGTGTTCATCAAAAGAGAGAGATATGATCTTATATATATAT 2403

QY 702 TyrAsnSerValPheMetGlnArgLeuLysIleAsnIleLeuGlnIleAlaSerThrAla 721

Db 2404 TATAACTGAGTCTTCATCAAGAGACAGACAGACAGACAGACAGACAGACAGACAG 2463

QY 722 ProProThrLeuSerPheIleProHisIlePheArgSerProTyrLysPheProSerSer 741

Db 2464 CCCCCACTACCTTGACCAATACCTCAGATTCCTCCAGAGGCTTACAAAGTTCCTAG 2523

QY 742 ProLeuArgIleProGlyGlyAsnIleTyrIleSerPheLeuLysSerProTyrIleVal 761

Db 2524 CCGTTACGATTCCTGAGAGAGAGATATATATATTCATCCCTTAAGAGTTCATATATA 2583

QY 762 SerGluGlyLeuThrProIleThrIleLysMetThrProArgSerArgIleLeuValSerIle 781

Db 2584 TCAGAAATTCGCAATATCAAAATAAATAAGATCCCAAGATCAAGAAATCTAGATATAT 2643

QY 782 GlyGluSerPheGlyThrSerGluLysPheGlyLysIleAsnGlnMetValCysAsnSer 801

Db 2644 CATTCAATCTCCGGATCTTCAGCAATCTTCACAAATCTTCACAAATAAATCAGATG 2703

QY 802 AspArgValIleLysArgSerAlaGlyLysSerAsnGlnProGlySerLeuLysLysLeu 821

Db 2704 GACCGTCTCTTAAATATATGTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2763

QY 822 ArgPheAspIleGluGlySerAspGlyAlaAspCysLysLysHisIleThrGlnSer 841

Db 2764 CCGTTTCATATTCAGAGATTCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2823

QY 842 LysPheThrIleLysLeuAlaGlyMetThrSerThrArgThrArgMetGlnLysSerIys 861

Db 2824 AAATTTCAGCAGAAATCTCCAGAAATCAGATCTACACCAACCAATCCAAAGACAGAA 2883

QY 862 MetAsnAspSerMetAspThrSerAsnLysGluLys 874

Db 2884 ATGAT 2922

RESULT 13

HUMBERS

LOCUS

DEFINITION Human retinoblastoma susceptibility mRNA, complete cds.

ACCESSION M15400

VERSION M15400.1 GI:190958

KEYWORDS retinoblastoma susceptibility.

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 243 to 4740)

AUTHORS Lee, W.H., Hookstein, R., Hong, P., Young, L.J., Shaw, J.Y., and Lee, J.Y.

TITLE Human retinoblastoma susceptibility gene: cloning, identification, and sequence

JOURNAL Science 235 (4794), 1394-1399 (1987)

MEDLINE 87149066

PUBMED 3823889

REFERENCE 2 (bases 1 to 480)

AUTHORS Lee, W.H., Hookstein, R., Young, L.J., Lin, C.J., Rosenfeld, M.J., and Lee, W.H.

TITLE Molecular mechanism of retinoblastoma gene inactivation in retinoblastoma cell line Y79

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (16), 6017-6021 (1988)

MEDLINE 88320373

PUBMED 3413073

COMMENT Draft entry and computer-readable copy of sequence in [1] kindly provided by F. Hookstein, 27 Apr-1987.

AK072041 AR072041 3332 bp DNA Linear PAT 18 FEB 2000
 DESCRIPTION Sequence 1 from patent US 5912236.
 ACCESSION AR072031
 VERSION AR072031.1 01:7222919
 KEYWORDS UNKNOWN.
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 3232)
 AUTHORS Xu, H., J., Ho, S.-X., and Benedict, W.F.
 TITLE Broad-spectrum tumor suppressor genes gene products and methods for tumor suppressor gene therapy
 JOURNAL Patent: US 5912236 A 15 JUN 1999.
 FEATURES
 location/Qualifiers
 1..3232
 /organism="unknown"
 BASE COUNT 1086 a 507 c 545 g 983 t
 ORIGIN

Alignment Scores:
 Prod. No.: 7,350-305 Length: 3232
 Score: 4213.00 Matches: 820
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 93.54% Indels: 0
 Gaps: 0

US 09-026-459A-31 (1-874) x AR072031 (1-3232)

QY 55 AspLeuAspGluMetSerPheThrGluLeuGlnLysAsnIleGluIleSerVal 74
 DB 7 GAGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 66
 QY 75 HistyPhePheAsnLeuLeuLysGluIleAspThrSerThrLysValAspAsnAlaMet 94
 DB 67 CATAAATCTTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTTACTT 126
 QY 95 SerArgLeuLeuLysLysTyrAspValLeuPheAlaLeuPheSerLysGluGluArgThr 114
 DB 127 TCAAGACATCTTCAAGACATCTTCAAGACATCTTCAAGACATCTTCAAGACATCT 186
 QY 115 CysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 134
 DB 187 TGTGAAATTAATAATTTGACACAACTCAGGATTCGATATCTACTGAAATTAATA 246
 QY 145 LeuValLeuLysValSerTrpIleThrPheLeuLeuAlaLysGlyGluValLeuGlnMet 154
 DB 147 TTGGTGATAAAGTTCTTGGATGATGATGATGATGATGATGATGATGATGATGAT 306
 QY 155 GluAspAspLeuValIleSerPheGlnLeuMetLeuCysValLeuAspTyrThrIleLys 174
 DB 407 GAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 366
 QY 175 LeuSerProMetLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 194
 DB 407 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
 QY 195 SerProArgThrProArgArgGlyGlnAsnArgSerAlaArgIleAlaLysGlnLeuGlu 214
 DB 427 TCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 486
 QY 215 AsnAspThrArgIleIleIleGluValLeuCysLysGluHisGlyCysAsnIleAspGluVal 234
 DB 487 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 546
 QY 245 LysAsnValTyrPheLysAsnPheIleProPheMetAsnSerLeuGlyLeuValThrSer 254
 DB 547 AAAAAAGTTATTCAAAAATTTTATACCTTTTATCAATTTCTTGACATTCATCA 606
 QY 265 AsnGlyLeuProGluValGluGluAsnLeuSerLysArgTyrGluLeuIleLysLysAsn 274
 DB 407 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 666

QY 275 LysAspLeuAspAlaValGluValLeuAspLysThrLeuGlnThrAspSerIle 294
 DB 667 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
 QY 295 AspSerPheGluThrGlnArgThrProArgLysSerAsnLeuAspGluGluValAsnVal 314
 DB 727 GATAGTTTAAAACATGATGATGATGATGATGATGATGATGATGATGATGAT 786
 QY 315 IleProProHisThrProValArgThrValMetAsnThrIleGlnGlnLeuMetMetIle 334
 DB 787 ATTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 846
 QY 335 LeuAsnSerAlaSerAspGlnProSerGluAsnLeuLeuSerTyrPheAsnAspCysThr 354
 DB 847 TTAATTCAGCAAGTGAATCAACCTTCAGAAAATCTGATATCCCTATTTTAACTGCA 906
 QY 365 ValAspProLysGluSerIleLeuLysArgValLysAspIleGlyTyrIlePheLysGlu 374
 DB 907 GTGAATGCAAAAAGATATAGTGAAGAAGATGAGATATAGATATATATTAAGAG 966
 QY 375 LysPheAlaLysAlaValGlyGlyGlyGlyCysValIleIleLysSerPheGluArg 394
 DB 967 AAATTTGCTAAAGCTGTGGAGATGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026
 QY 395 GlyValArgLeuTyrTyrArgValMetGluSerMetLeuLysSerGluGluLeuArgLeu 414
 DB 1027 GAGTTCGCTGCTATATACCCAGTAATGCAATCCATGCTTAATATCAAGAAACCA 1086
 QY 415 SerIleGlnAsnPheSerLysLeuLeuAsnAspAsnIlePheHisMetSerLeuLeuAla 434
 DB 1087 TCAATTCAAAATTTATGCAAAATTTCTGAATGATGATGATGATGATGATGATGAT 1146
 QY 435 CysAlaLeuGluValValMetAlaThrTyrSerArgSerThrSerGluAsnLeuAspSer 454
 DB 1147 TGAGTGTGTTGAGTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 1206
 QY 455 GlyThrAspLeuSerPheProTrpIleLeuAsnValLeuAsnLeuLysAlaPheAspPhe 474
 DB 1207 GGAACAGATTGCTGCTCCCATGCAATCTCTCAATGCTCTTAATTTAAAGCCCTT 1266
 QY 475 TyrLysValIleGluSerPheIleLysAlaGluGluValLeuThrArgGluMetIleLys 494
 DB 1267 TACAAATGATGCAAAATTTTATTAAGATGATGATGATGATGATGATGATGATGAT 1426
 QY 495 HisLeuGluArgCysGluHisArgIleMetGluSerLeuAlaIlePheSerAspSerPro 514
 DB 1327 CATTTAAGATGCTGAACATCGAATCATGGAATCCCTTCATGCTCTCATATTCAC 1386
 QY 515 LeuPheAspLeuIleLysGlnSerLysAspArgGluGlyProThrAspHisLeuGluSer 534
 DB 1387 TTATTTGATCTTATTAACCAATCAAGAGCCGAGAGCCGAGAGCCGAGAGCCGAG 1446
 QY 535 AlacysProLeuAsnLeuProLeuGlnAsnAsnHisThrAlaAlaAspMetTyrLeuSer 554
 DB 1447 GCTTGTGCTCTTAATCTTCTGAGATTAATCACTGACAGATATATGATCTTCT 1506
 QY 555 ProValArgSerProLysLysCysLysSerThrThrArgValAsnSerThrAlaAsnAla 574
 DB 1507 CTTGTAAATCTGCAAAAAGAGGTTCAAGATGAGGTTGATTAATTTCAATTTCA 1566
 QY 575 GluThrGlnAlaThrSerAlaPheGlnThrGluLysProLeuLysSerThrSerLeuSer 594
 DB 1567 GAGATCAAGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1626
 QY 595 LeuPheTyrLysValTyrArgLeuAlaTyrLeuArgLeuAsnThrLeuLysGluArg 614
 DB 1627 CTGTTTTAATAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686
 QY 615 LeuLeuSerGluHisProGluLeuGluHisIleIleTrpThrLeuPheGlnHisThrLeu 634
 DB 1687 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1746

```

57 636 GluAsnGluTyrIleLeuMetArgAspArgHisLeuAspGlnIleMetMetCysSerMet 654
14 1747 CACAATGACTATACAACTCAGACACAGGCAATGGAGCAAAATATATCATGTGTTCGACG 1806
57 655 TGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 674
14 1807 TATGGTATATGTAAGAGGAAATATAGATTAATAATCAAAATCATGTAAGAGCATAC 1866
57 675 LysAspLeuProHisAlaValGlnIleThrPheTyrSerValIleGluTyrGlnGluGlu 694
14 1867 AAGGATCTCCCTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1926
57 695 TyrAspSerIleIleValPheTyrAsnSerValPheMetGlnArgLeuLysThrAsnIle 714
14 1927 TATGATCTATATATGATATCTATTAACCTGGCTCTCTCTCTCTCTCTCTCTCTCTCT 1986
57 715 LeuGlnTyrTAlaSerThrArgProThrLeuSerProIleProHisIleProArgSer 734
14 1987 TTGCACTATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2046
57 735 ProTyrLysPheProSerSerProLeuArgIleProGlnGlnGlnGlnGlnGlnGln 754
14 2047 CATTAAAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2106
57 755 LeuLysSerProTyrLysIleSerGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 774
14 2107 CTGCAAGAGCTGATATAAAATCTTCAAGAGCTGCTGGCAATACCAATAAAAAGACCT 2166
57 775 SerArgIleLeuValSerIleGlnSerPheThrSerGlnLysPheGlnLysIle 794
14 2167 TAAATAATCTTACTATCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2226
57 795 AsnGlnMetValCysAsnSerAspArgValLeuLysArgSerAlaGlnGlnGlnSer 814
14 2227 AATACATGCTATGTAATAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2286
57 815 ProLysProLeuLysIleArgPheAspIleGlnGlnGlnGlnGlnGlnGlnGlnGln 834
14 2287 CTTAAACCTACTCAAAAATACGTTTTCATATGAAAGATCAGATGAGCATGGAAGT 2346
57 835 LysHisLeuProGlnTyrGlnSerLysPheGlnGlnLysLeuAlaGluMetThrSerThrArg 854
14 2447 AAACATCTCTCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
57 855 ThrArgMetGlnLysIleLysMetAsnAspSerMetAspThrSerAsnLysGlnGlnLys 874
14 2407 AATCAATCTCAAAAGTCAAAAATGAAATGATATGGAATACCTCTCAAAACAGCAAGAAA 2466
```

Search completed: January 18, 2003, 22:15:50
Job time: 4000.04 secs




```

1879 GCAACAAATGATACCTTGAAGATCTGTGTCCTTAATCTTCCTCCAGAAATAATCAC 1938
QY 542 ThrAlaAlaAspMetLysLeuSerProValArgSerProLysLysLysLysLysLys 561
1879 ACTGACGAGATATATCTTCTGCTGAGATCTCTCAAGAGAAAGAGGTTCAGATCG 1998
QY 562 ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGluThrGlnLys 581
1879 CAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 582 ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrArgLeuAlaTyrLeu 601
1879 CAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 602 ArgLeuAsnThrLeuGlySerLeuArgLeuLeuSerGluHisProGluGluHisLeuLeu 621
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 642 TrpThrLeuPheGluHisThrLeuGlnAsnGluTyrGluLeuMetArgAspArgHisLeu 641
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 642 AspGluHisMetMetCysSerMetTyrGluLeuGlySerValLysAsnLysAspLeuLys 661
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 662 PheGlySerLeuValThrAlaTyrLysAspGluProHisAlaValGluGluThrPheLys 681
1879 TTTCAAAATCATGTAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 2458
QY 682 ArgValLeuLeuLysGluGluGluTyrAspSerLeuLeuValPheTyrAsnSerValPhe 701
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 702 MetGluArgLeuLysThrAsnLeuLeuGluTyrAlaSerThrArgProThrLeuSer 721
1879 ATGAGAGAGCTGAGAAATATTTTGGAGTATGCTTCCACAGAGGCGGCTACCTTGCTCA 2478
QY 722 ProLeuHisThrProArgSerProLysPheProLysSerProLysSerProLysSerPro 741
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 742 GlyGlyAsnLeuTyrThrSerProLeuLysSerProLysLysSerGluGluLeuPro 761
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 762 ThrProThrLysMetThrProArgSerArgLeuValSerLeuGlyLysSerPheGly 781
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 782 ThrSerGluLysPheGluLysThrAsnGluMetValCysAsnSerAspArgValLeuLys 801
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 802 ArgValLeuLeuLysSerProLeuLysSerProLeuLysSerProLeuLysSerPro 821
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 822 GlySerAspGluAlaAspGlySerLysGluSerProGlyLysSerPheGluGlnLys 841
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 842 LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetAspAspSerMet 861
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 862 AspThrSerAsnLysGluGluLys 889
1879 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
QY 889 GAGTCTAAATCTTCTTCTGAAATCTCTCTCTGAGTACGCAAGCAATTAACATATCATC 2178
RESULT 2
AR09189

```

```

LOCUS AR09189
DEFINITION Sequence 3 from patent US 6074850.
ACCESSION AR09189
VERSION AR09189.1 GI:12807446
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2994)
AUTHORS Autelman, D., Gregory, K.J., and Willis, K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 4 14 JUN 2000;
FEATURES
    Location/Qualifiers
        source
            1..2994
            /organism="unknown"
BASE COUNT 974 a 618 c 593 q 809 t
ORIGIN
Alignment Scores:
    Pred. No.: 9,84e-312 Length: 2994
    Score: 4454.50 Matches: 867
    Percent Similarity: 93.53% Conservat: 1
    Best Local Similarity: 93.43% Mismatches: 1
    Query Match: 98.94% Indels: 54
    DB: 6 Gaps: 1
US 09-026-459A-47 (1-869) x AR09189 (1-2994)
QY 1 MetProLysThrProArgLysThrAlaAlaThrAlaAlaAlaAlaAlaAlaPro 20
DB 149 ATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 198
QY 21 ProAlaProProProProProProGluGluAspProGluGluAspProGluGlu 40
DB 199 CCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 258
QY 41 AspLeuProLeuValArgLeuGluPheGluGluThrGluGluProAspPheThrAlaLeu 60
DB 259 GACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 118
QY 61 CysGlnLysLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLys 80
DB 319 TGTGAGAAATTAAGATACATCATCATCATCATCATCATCATCATCATCATCATTA 478
QY 81 ValSerSerValAspGlyValLeuGlyLysLysLysLysLysLysLysLysLys 100
DB 379 GCTTCATCTGCTGGATGGAGATATGGGAGGTTATATCAAAAGAGAGAGAGAGAG 448
QY 101 IleCysIlePheIleAlaAlaValAspLeuAspGluMetSerPheThrPheThrGluLeu 120
DB 439 ATCTGATCTTTATTTGAGAGAGTTGAGCGTATGATGATGATGATGATGATGATG 498
QY 121 GlnLysAsnIleGluIleSerValHisLysPhePheAsnLeuLeuLysGluLeuAsp 140
DB 499 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 558
QY 141 SerThrLysValAspAsnAlaMetSerArgLeuLeuLysLysLysLysLysLys 160
DB 559 ACTACCAAGTGTGATTAATGCTATGCTGAGAGATGCTGAGAGAGAGAGAGAGAG 618
QY 161 LeuPheSerLysLeuGluArgThrCysGluLeuLeuLysLeuThrGluProSerSer 180
DB 619 CTTCTTCAGCAAAATTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 678
QY 181 Met 181
DB 679 ATATCTACTGAAATATAATTTTGTGATTGTGTGTGTGTGTGTGTGTGTGTGT 748
QY 181 181
DB 739 GCTAAAGCGGAGATTTACAAATGTAACATGATGATGATGATGATGATGATGATG 798
QY 181 181

```

```

Db 799 GAGGAGTATGACCTATTTATTAACCTCTCACCTCCCATGTTGCTCAAGAACCATATAAA 858
QY ValAlaValIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 201
Db 859 ACAGCTGTTATACCAATATATGGTTACCTCCGACACCCAGCGAGGTGAGAACAGAGCT 918
QY AlaArgAlaValIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 221
Db 919 GCACGATATAGCAAAATAGTAAATATATAGTAAATATATAGTAAATATATAGTAAATAT 978
QY HisGlnGlyAsnIleAspGluValIleAspGluValIleAspGluValIleAspGluVal 241
Db 979 CATCAATATATATATATATATATATATATATATATATATATATATATATATATATAT 1038
QY AsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGluAsnLeuSerLysArg 261
Db 1039 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1098
QY TyrGluGluIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 281
Db 1099 TACGAGAAATTTATATATATATATATATATATATATATATATATATATATATATAT 1158
QY LysThrLeuGluIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 301
Db 1159 AAAACCTCTTACAGCTGATCTATAGACGTTTGAACACACACAGACACACCAAAAGT 1218
QY AsnLeuAspGluGluValIleProIleAsnGlySerProArgThrProArgArgGlyGln 321
Db 1219 AACCTTCAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1278
QY ThrIleGlnGluIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 341
Db 1279 ACTATGCAACAAATATATATATATATATATATATATATATATATATATATATATAT 1338
QY IleSerIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 361
Db 1339 ATTTCCCTATATATATATATATATATATATATATATATATATATATATATATAT 1398
QY AspIleGlyTyrIleProIleAsnGlySerProArgThrProArgArgGlyGlnAsnArgSer 381
Db 1399 GATATAGATATATATATATATATATATATATATATATATATATATATATATATAT 1458
QY IleCysSerGlnArgTyrIleProIleAsnGlySerProArgThrProArgArgGlyGln 401
Db 1459 ATTGATATATATATATATATATATATATATATATATATATATATATATATATAT 1518
QY LeuIleGlySerGlnArgTyrIleProIleAsnGlySerProArgThrProArgArgGlyGln 421
Db 1519 GTTAAATATATATATATATATATATATATATATATATATATATATATATATAT 1578
QY IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrTyrSerArg 441
Db 1579 ATTTCCCTATATATATATATATATATATATATATATATATATATATATATATAT 1638
QY SerIleSerGlnArgTyrIleProIleAsnGlySerProArgThrProArgArgGlyGln 461
Db 1639 AGTAAATATATATATATATATATATATATATATATATATATATATATATATAT 1698
QY LeuAsnLeuIleAlaPheAspGlyTyrValIleGluSerPheIleLysAlaGluGly 481
Db 1699 GTTAAATATATATATATATATATATATATATATATATATATATATATATATAT 1758
QY AsnLeuThrArgGluMetIleLysHisLeuGluAlaCysGluIleHisArgIleMetGluSer 501
Db 1759 AACTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1818
QY LeuAlaIlePheSerAspSerProIleAsnGlySerProArgThrProArgArgGlyGln 521
Db 1819 GTTCAATATATATATATATATATATATATATATATATATATATATATATATAT 1878
QY GlyProThrAspGlySerGluAlaCysProIleAsnGlySerProArgThrProArgArgGly 541

```

RESULT 3

AR207831

LOCUS

AP207831

2994 bp

DNA

Linear

PAT 20-JUN-2002

```

Db 1879 GCACCAATCATCAACCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1938
QY ThrAlaAlaAspMetIleLeuSerProValAlaSerProLysLysLysLysLysLysLys 561
Db 1939 AATATATATATATATATATATATATATATATATATATATATATATATATATATAT 1998
QY ArgValAsnSerThrAlaAsnAlaGluThrGlnAlaThrSerAlaPheGlnThrGlnLys 581
Db 1999 GGTATATATATATATATATATATATATATATATATATATATATATATATATATAT 2058
QY ProLeuLysSerThrSerLeuSerLeuPheTyrLysLysValTyrArgLeuAlaTyrLeu 601
Db 2059 CCAATGAAATCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2118
QY ArgLeuAsnThrLeuCysGluAlaLeuLeuSerGluIleProGluLeuGluIleIleIle 621
Db 2119 CGGTAAATATATATATATATATATATATATATATATATATATATATATATATAT 2178
QY TrpThrLeuPheGlnHisThrLeuGlnAsnGluTyrGluLeuMetArgAspArgHisLeu 641
Db 2179 TGAATATATATATATATATATATATATATATATATATATATATATATATATATAT 2238
QY AspGlnIleMetLysSerMetIleGlyIleCysLysValLysAsnIleAspLeuLeu 661
Db 2239 GACCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 2298
QY PheLysIleIleValThrAlaTyrLysAspLeuProHisAlaValGlnGlnThrPheLys 681
Db 2299 TTTCAAAATATATATATATATATATATATATATATATATATATATATATATATAT 2358
QY ArgValLeuIleLysGluGluGluTyrAspSerIleIleValPheTyrAsnSerValPhe 701
Db 2359 CGTCTTTTATATATATATATATATATATATATATATATATATATATATATATAT 2418
QY MetGlnArgLeuLysThrAsnIleLeuGlnTyrAlaSerThrArgProGluLeuSer 721
Db 2419 ATCCACAGACTGAAACAAATATATATATATATATATATATATATATATATATATATAT 2478
QY ProIleProHisIleProArgSerProTyrLysPheProSerSerProLeuArgIlePro 741
Db 2479 CCAATATATATATATATATATATATATATATATATATATATATATATATATAT 2538
QY GlyCysAsnIleCysThrProLeuLysSerProTyrLysIleSerGluGlyLeuPro 761
Db 2539 GATATATATATATATATATATATATATATATATATATATATATATATATATATAT 2598
QY ThrProThrLysMetThrProArgSerAlaGluLeuValSerIleCysGluSerPheGly 781
Db 2599 ACACCAACAAAAAGACTCCACATCAACAAATCTTACTATCAATTCGTCGCAATCCGG 2658
QY ThrSerGluLysPheGlnLysIleAsnGlnMetValCysAsnSerAspArgValLeuLys 801
Db 2659 ACTCTGAGAAATTCGAGAAATATATATATATATATATATATATATATATATATATAT 2718
QY ArgSerAlaCysGlySerAsnProPheLysProLeuLysLysLeuArgPheAspIleCys 821
Db 2719 AGAAATATATATATATATATATATATATATATATATATATATATATATATATAT 2778
QY GlySerAspGluAlaAspGlySerLysHisLeuProGlyGluSerLysPheGluGlnLys 841
Db 2779 GCATCATATCAACACATGAGAAATATATATATATATATATATATATATATATATATAT 2838
QY LeuAlaGluMetThrSerThrArgThrArgMetGlnLysGlnLysMetLysAsnSerMet 861
Db 2839 CTGAGTAAATATATATATATATATATATATATATATATATATATATATATATAT 2898
QY AspThrSerAsnLysGluGlnLys 869
Db 2899 GATATATATATATATATATATATATATATATATATATATATATATATATATAT 2922

```

DEFINITION Sequence from patient US 6379927.
 ACCESSION AR207841.1 GI:21607487
 VERSION AR207841.1 GI:21607487
 KEYWORDS unknown.
 ORGANISM unknown.
 classification
 REFERENCE 1 (bases 1 to 2994)
 AUTHORS Antelman,D., Gregory,R.J., and Willis,K.N.
 TITLE Notch1/Notch2 fusion proteins
 JOURNAL Patent: US 6,379,927 A 3 26 Apr 2002.
 FEATURES
 Location/Qualifiers
 1..2994
 /organism="unknown"
 BASE COUNT 974 a 618 c 594 g 809 t
 ORIGIN
 Alignment Statistics:
 Prod. No.: 9,846,412 Length: 2994
 Source: 44,94,50 Matches: 867
 Percent Similarity: 93.54%
 Best Local Similarity: 93.43%
 Query Match: 98.94%
 Indels: 59
 Gaps: 1
 US 09 026 459A 47 (1 869) x AR207841 (1-2994)
 QY 1 MetProProGysThrProArLysThrAlaAlaThrAlaAlaAlaAlaAlaGluPro 20
 DB 149 ARGGGGGAAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 198
 QY 21 ProAlaProProProProProProGluGluAspProGluGluAspSerGlyProGlu 40
 DB 199 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 258
 QY 41 AspLeuProLeuValArgLeuGluPheGluGluThrGluGluProAspPheThrAlaLeu 60
 DB 259 GAGTGGCTCTCTCAGCTTGAATTTTCAGGAAACAGAGAAAGCTGATTTTACGATTA 318
 QY 61 CysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 80
 DB 319 TGTCAGAAATTAAGAGATAGATGATGATGATGATGATGATGATGATGATGATGAT 378
 QY 81 ValSerSerValAspGlyValLeuGlyValLeuGlyValLeuGlyValLeuGlyVal 100
 DB 379 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 438
 QY 101 ThrCysThrPheThrAlaAlaValAspLeuAspGluMetSerPheThrPheThrGluLeu 120
 DB 439 ATCTTATTTTATTCAGAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 498
 QY 121 GluLysAsnGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 140
 DB 499 CAGAAAAAATAGAAATAGATGATGATGATGATGATGATGATGATGATGATGATG 558
 QY 141 SerThrValAspAsnAlaMetSerArgLeuLeuLysLysTyrAspValLeuPheAla 160
 DB 559 AGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 618
 QY 161 LeuPheSerLysLeuGluAlaThrCysGluLeuLeuLeuLeuLeuLeuLeuLeuLeu 180
 DB 619 CTCTTCAGCAAAATGGAAAGAGATGAGGATTAATATATATATATATATATATAT 678
 QY 181 Met
 DB 679 ATATGTATTAATAAATATGTGATTGGTGTAAAGTTTCTTGGATCAATTTTATTA 738
 QY 181
 DB 739 GTTAAAGAGAGATTAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 798
 QY 181

DB 799 TGTGTCCTTGAATTTATTTTAAATCTCTATCTCTCTCTCTCTCTCTCTCTCTCT 858
 QY 182 ValAlaValIleProIleAsnGlySerThrArgThrProAlaAlaGlyGluAsnAlaSer 201
 DB 859 ACAGTCTGTTATACCATTAATGGGTCACCTCGAAACATCTACGTCACCTCTCTCTCT 918
 QY 202 AlaArgIleAlaLysGluLeuGluAsnAspThrArgIleIleGluValLeuLysGlu 221
 DB 919 GCACGGATAGCAAAACACATAGAAAATCATACAAAGAAATATCTAAGCTCTCTCTCTCT 978
 QY 222 HisSerLysAsnGluValLysAsnValLysLysAsnGluValLysAsnGluValLys 241
 DB 979 CATGAATGTATATAGATGAGGTGAAAAATGTTTATTTAAAAATTTTATATTTTATG 1038
 QY 242 AsnSerLeuGlyLeuValIleSerAsnGlyLeuGluValIleGluAsnLeuSerLysArg 261
 DB 1039 AATCTCTTGGACCTGTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1098
 QY 262 TyrGluGluIleTyrLeuLysAsnGlyLysLeuAlaAlaAlaAlaAlaAlaAlaAla 281
 DB 1099 TACGAGAGAAATTAATCTTAAAAATTAAGATCTACATCTCTCTCTCTCTCTCTCTCT 1158
 QY 282 LysThrLeuGluThrAspSerIleAspSerPheGluThrGluAlaGluProAlaLysSer 301
 DB 1159 AAAATCT 1218
 QY 302 AsnLeuAspGluGluValAsnValIleProProHisThrProValAlaThrValMetAsn 321
 DB 1219 AACCTTCAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1278
 QY 322 ThrIleGluGluLeuMetIleLeuAsnSerAlaSerAspIleProSerGluAsnLeu 341
 DB 1279 ACTATCCAACTAATTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1338
 QY 342 IleSerTyrPheAsnGlyThrValAspProLysLysLysLysLysLysLysLysLys 361
 DB 1339 ATTTCTCTATTTTACCACTGACAGTGAATGCAAAAGAAATATCTGAAAGAGTGAG 1398
 QY 362 AspIleGlyTyrIlePheLysGluLysPheAlaLysAlaValGlyIleGlyLysValGlu 381
 DB 1399 GATATAGATACATCTCTTAAAGAGAAATTTCTTAAAGCTCTCTCTCTCTCTCTCTCT 1458
 QY 382 IleGlySerGluArgTyrLysLeuGlyValArgLeuTyrTyrAlaValMetGluSerMet 401
 DB 1459 ATTGGATCAACAGCATACAAACCTTGGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1518
 QY 402 LeuLysSerGluGluGluArgLysSerIleGluLysAspSerLysLysLysLysLysLys 421
 DB 1519 CTTAAATCAGAAAGAGAGGATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1578
 QY 422 IlePheHisMetSerLeuLeuAlaCysAlaLeuGluValValMetAlaThrLysSerArg 441
 DB 1579 ATTTTCTCATCT 1638
 QY 442 SerThrSerGluAsnLeuAspSerGlyThrAspLeuSerPheProThrIleLeuAsnVal 461
 DB 1639 ACTACATCT 1698
 QY 462 LeuAsnLeuLysAlaPheAspPheLysValIleGluSerPheIleLysAlaGluGly 481
 DB 1699 CTTAATTTAAAGGCTTTGATTTTAAAGATGATGAAATTTTATCTCTCTCTCTCTCT 1758
 QY 482 AsnLeuThrArgGluMetIleLysHisLeuGluArgLysGluHisArgIleMetGluSer 501
 DB 1759 AACCTTCAACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1818
 QY 502 LeuAlaTrpLeuSerAspSerProLeuPheAspLeuLysLysLysLysLysLysLys 521
 DB 1819 CTTCATGCT 1878
 QY 522 GlyProThrAspHisLeuGluSerAlaCysProLeuAsnLeuProLeuGluAsnAsnHis 541
 DB 1879 GGAGCAACTGATCATCTCTGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1948


```

642 LeuAlaSerMetThrSerThrArgMetGlnLysGlnLysMetAsnAspSerMet 861
|||||
181 Met----- 181
679 ATATCTACTGAAATATAATCTGCAATGGTCTAAAGATTTCTTGGATCACAATTTTATTA 748
|||||
181 ----- 181
739 GTTAAACGGCAAGATTTACAAATGGCAACATGCTGCTCATTTTCACATTAATCCCA 798
|||||
181 ----- 181
759 TGGTCTTGATTAATTTTATTAATCTGCAATGGTCTAAAGATTTCTTGGATCACAATTA 858
|||||
182 ValAlaValIleProIleAsnGlySerProArgIleProArgArgGlyGlnAsnArgSer 201
|||||
859 ACAGTGTGTATACCAATTAATGTTTACCTGCAACCCAGGCGAGCTCAGAACAGAGG 918
|||||
202 AlaArgIleAlaLysGlnLeuGluAspThrArgIleGluValLeuGlyGlySer 221
|||||
919 GTATGCAATAGTAAACCAACATGAAAAAGATTAATGCAAGTCTCTCTAAACAA 978
|||||
222 HisGluCysAsnIleAspGluValLysAsnValTyrIleLysAsnIleProIleMet 241
|||||
979 CATGATGTATATATATATGAGTGAATAATGTTTATTTCAAAAATTTTATATAT 1038
|||||
242 AsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGlnLeuSerLysArg 261
|||||
1039 AATCTCTTGCACATTTGAACTAATGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 1098
|||||
262 TyrGluGluIleTyrLeuLysAsnLysAspLeuAspAlaValLeuPheLeuAspHis 281
|||||
1099 TACCAACAAATTAATCTTAAATAAAGATCTAGATGCAAGATTAATTTTGGATCAT 1158
|||||
282 LysThrLeuGluThrAspSerIleAspSerPheGluThrGlnArgThrProArgLys 301
|||||
1159 AAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1218
|||||
302 AsnLeuAspGluGluValAlaValIleProProHisThrProValArgThrValMetAsn 321
|||||
1219 AACCTTCATCAAGAGGTGAATGTAATTCCTCCACACACATCCACTTAGGACTTCATGAAC 1278
|||||
322 ThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 341
|||||
1279 ACTATCCAAATTAATGATGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1338
|||||
342 TCTSerTyrPheAsnAspCysThrValAsnProGlySerLeuSerIleLeuLysArgVal 361
|||||
1339 ATTTCTTATTTTAACTATCTATAGTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1398
|||||
362 AspIleGlyTyrIleIleLysGlyPheAlaLysAlaValIleGlyGlyCysValGlu 381
|||||
1399 GATATAGCAATACATTTTAAAGAGAAATTTGCTAAAGCTGGGACAGGCTTCTGTCGAA 1458
|||||
382 TleLysSerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArgValMetGlySerMet 401
|||||
1459 ATTGGATCAGAGGATACAACTTGGAGTGGCTTATTTACCGAGTAAAGAAATCCCA 1518
|||||
402 LeuLysSerGluGluIleArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAsp 421
|||||
1519 CTTAATATCAACAAAGAAAGATTAATCCATCAAAATTTTAAAGAAATTTTAAAGAA 1578
|||||
422 TlePheHisMetSerLeuLeuAlaCysAlaLeuGluValIleMetAlaThrTyrSerArg 441
|||||
1579 APTTTCATATGCTTTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638
|||||
442 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProTyrPheLeuAsnVal 461
|||||
1639 AGTACATCTCAGAAATCTTATCTTCCAGACAGATTTGCTTTCCCATGGATTCGATG 1698
|||||
462 LeuAsnLeuLysAlaPheAspPheTyrLysValIleGlnSerPheIleLysAlaGluGly 481
|||||
1699 CTTAATTTAAAGAGCTTTGATTTTTTAAAGAGCTTTTAAAGAGCTTTTAAAGAGAG 1758
|||||

```

642 LeuAlaSerMetThrSerThrArgMetGlnLysGlnLysMetAsnAspSerMet 861

181 Met----- 181

679 ATATCTACTGAAATATAATCTGCAATGGTCTAAAGATTTCTTGGATCACAATTTTATTA 748

181 ----- 181

739 GTTAAACGGCAAGATTTACAAATGGCAACATGCTGCTCATTTTCACATTAATCCCA 798

181 ----- 181

759 TGGTCTTGATTAATTTTATTAATCTGCAATGGTCTAAAGATTTCTTGGATCACAATTA 858

182 ValAlaValIleProIleAsnGlySerProArgIleProArgArgGlyGlnAsnArgSer 201

859 ACAGTGTGTATACCAATTAATGTTTACCTGCAACCCAGGCGAGCTCAGAACAGAGG 918

202 AlaArgIleAlaLysGlnLeuGluAspThrArgIleGluValLeuGlyGlySer 221

919 GTATGCAATAGTAAACCAACATGAAAAAGATTAATGCAAGTCTCTCTAAACAA 978

222 HisGluCysAsnIleAspGluValLysAsnValTyrIleLysAsnIleProIleMet 241

979 CATGATGTATATATATATGAGTGAATAATGTTTATTTCAAAAATTTTATATAT 1038

242 AsnSerLeuGlyLeuValThrSerAsnGlyLeuProGluValGlnLeuSerLysArg 261

1039 AATCTCTTGCACATTTGAACTAATGCTTCCAGCTTCCAGCTTCCAGCTTCCAGCT 1098

262 TyrGluGluIleTyrLeuLysAsnLysAspLeuAspAlaValLeuPheLeuAspHis 281

1099 TACCAACAAATTAATCTTAAATAAAGATCTAGATGCAAGATTAATTTTGGATCAT 1158

282 LysThrLeuGluThrAspSerIleAspSerPheGluThrGlnArgThrProArgLys 301

1159 AAAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1218

302 AsnLeuAspGluGluValAlaValIleProProHisThrProValArgThrValMetAsn 321

1219 AACCTTCATCAAGAGGTGAATGTAATTCCTCCACACACATCCACTTAGGACTTCATGAAC 1278

322 ThrIleGlnGlnLeuMetMetIleLeuAsnSerAlaSerAspGlnProSerGluAsnLeu 341

1279 ACTATCCAAATTAATGATGATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1338

342 TCTSerTyrPheAsnAspCysThrValAsnProGlySerLeuSerIleLeuLysArgVal 361

1339 ATTTCTTATTTTAACTATCTATAGTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 1398

362 AspIleGlyTyrIleIleLysGlyPheAlaLysAlaValIleGlyGlyCysValGlu 381

1399 GATATAGCAATACATTTTAAAGAGAAATTTGCTAAAGCTGGGACAGGCTTCTGTCGAA 1458

382 TleLysSerGlnArgTyrLysLeuGlyValArgLeuTyrTyrArgValMetGlySerMet 401

1459 ATTGGATCAGAGGATACAACTTGGAGTGGCTTATTTACCGAGTAAAGAAATCCCA 1518

402 LeuLysSerGluGluIleArgLeuSerIleGlnAsnPheSerLysLeuLeuAsnAsp 421

1519 CTTAATATCAACAAAGAAAGATTAATCCATCAAAATTTTAAAGAAATTTTAAAGAA 1578

422 TlePheHisMetSerLeuLeuAlaCysAlaLeuGluValIleMetAlaThrTyrSerArg 441

1579 APTTTCATATGCTTTTATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1638

442 SerThrSerGlnAsnLeuAspSerGlyThrAspLeuSerPheProTyrPheLeuAsnVal 461

1639 AGTACATCTCAGAAATCTTATCTTCCAGACAGATTTGCTTTCCCATGGATTCGATG 1698

462 LeuAsnLeuLysAlaPheAspPheTyrLysValIleGlnSerPheIleLysAlaGluGly 481

1699 CTTAATTTAAAGAGCTTTGATTTTTTAAAGAGCTTTTAAAGAGCTTTTAAAGAGAG 1758

